

```

1 CCATTCCAAA CAAGTCAGGA AAGCCTGCAC AGGACTGGAT AAATAATTAA
51 GAACAGAGTG TTCTGAACAT CAACACAAAG TGAAGAACC TTAAGCTGAA
101 GGTACAGTAT ATTATTTACA CTGAAGGGGC TTGTGTGTGG ACAAGAAAGC
151 GCTGACAGCT CAAATGGATC CCATGGAAC TGAAGATGTC AACATCGAAC
201 CAGATGATGA GAGCAGCAGT GGAGAAAGTG CTCCAGATAG CTACATCAGG
251 ATAGGAAATT CAGAAAAGGC AGCAATGAGC AGTCAATTTG CTAATGAAGA
301 CACTGAAAGT CAGAAATTCC TGACAAATGG ATTTTGGGG AAAAAGAAGC
351 TGGCAGATTA TGCTGATGAA CACCATCCCG GAACCACTTC CTTTGAATG
401 TCTTCATTTA ACCTGAGTAA TGCCATCATG GGCAGTGGGA TCCTGGGCTT
451 GTCCATATGCC ATGGCCTACA CAGGGGTCAT ACTTTTATA ATCATGCTGC
501 TTGCTGTGGC AATATTATCA CTGTATTTCAG TTCACCTTTT ATTA AAAACA
551 GCCAAGGAAG GAGGGTCTTT GATTTATGAA AAATTAGGAG AAAAGGCATT
601 TGGATGGCCG GGAAAAATTG GAGCTTTTGT TTCCATTACA ATGCAGAAC
651 TTGGAGCAAT GTCAAGCTAC CTCTTTATCA TTAATATGA ACTACCTGAA
701 GTAATCAGAG CATTCATGGG ACTTGAAGAA AATACTGGAG AATGGTACCT
751 CAATGGCAAC TACCTCATCA TATTTGTGTC TGTTGGAATT ATTCTCCAC
801 TTTGCTCCT TAAAAATTTA GGTATCTTG GCTATACCAG TGGATTTTCT
851 CTACCTGCA TGGTGT TTTT TGTTAGTGTG GTGATTTACA AGAAATTCCA
901 AATACCCTGC CCTCTACCTG TTTTGGATCA CAGTGTGGA AATCTGTCAT
951 TCAACAACAC GCTTCCAATG CATGTGGTAA TGTTACCAA CAACTCTGAG
1001 AGTTCTGATG TGAACCTCAT GATGGATTAC ACCCACCACA ATCCTGCAGG
1051 GCTGGATGAG AACCAGGCCA AGGGCTCTCT TCATGACAGT GGAGTAGAAT
1101 ATGAAGCTCA TAGTGATGAC AAGTGTGAAC CCAAATACTT TGTATTCAAC
1151 TCCCGGACGG CCTATGCAAT TCCTATCCTA GTATTTGCTT TTGTATGCCA
1201 CCCTGAGGTC CTTCCCATCT ACAGTGAAC TAAAGATCGG TCCCGGAGAA
1251 AAATGCAAAC GGTGTCAAAT ATTTCCATCA CGGGGATGCT TGTATGTAC
1301 CTGCTTGCCG CCTCTTTTGG TTACCTAACC TTCTATGGAG AAGTTGAAGA
1351 TGAATTACTT CATGCCTACA GCAAAGTGT TACATTAGAC ATCCCTCTTC
1401 TCATGGTTCG CCTGGCAGTC CTTGTGGCAG TAACACAAAC TGTGCCCATT
1451 GTCCTCTTCC CAATTCGTAC ATCAGTGATC AACTGTAT TTCCCAAACG
1501 ACCCTTCAGC TGGATACGAC ATTTCTGAT TGCAGCTGTG CTTATTGCAC
1551 TTAATAATGT TCTGGTCATC CTTGTGCCAA CTATAAAATA CATCTTCGGA
1601 TTCATAGGGG CTTCTTCTGC CACTATGCTG ATTTTATTC TTCCAGCAGT
1651 TTTTATCTT AACTTGTCA AGAAAGAAC TTTTAGGTCA CCCCCAAGG
1701 TCGGGGCTTT AATTTCTCTT GTGGTTGGAA TATTCCTCAT GATTGGAAGC
1751 ATGGCACTCA TTATAATTGA CTGGATTAT GATCTCCAA ATTCCAAGCA
1801 TCACTAACAC AAGGAAAAAT AC

```

**FEATURES:**

5'UTR: 1-163  
Start Codon: 164  
Stop Codon: 1805  
3'UTR: 1808

# **HOMOLOGOUS PROTEINS:**

## Top BLAST Hits:

	Score	E
CRA 145000039337444 /altid=gi 12017941 /def=gb AAG45335.1 AF295...	975	0.0
CRA 114000033649823 /altid=gi 10945621 /def=gb AAG24618.1 AF298...	597	e-169
CRA 160000003782430 /altid=gi 8677401 /def=gb AAF75589.2 AF1736...	591	e-168
CRA 148000002720069 /altid=gi 8248427 /def=gb AAF74195.1 AF2496...	587	e-166
CRA 87000000006802 /altid=gi 7243145 /def=dbj BAA92620.1  (AB03...	578	e-164
CRA 18000005069115 /altid=gi 5870893 /def=ref NP_006832.1  tran...	500	e-140
CRA 88000001154721 /altid=gi 7406950 /def=gb AAF61849.1 AF15985...	496	e-139
CRA 66000019404613 /altid=gi 9506837 /def=ref NP_061849.1  amin...	495	e-139
CRA 100000004435450 /altid=gi 8926332 /def=gb AAF81797.1 AF2730...	492	e-138
CRA 335001098689635 /altid=gi 11434147 /def=ref XP_006635.1  hy...	480	e-134

## EST:

gi 10934204 /dataset=dbest /taxon=96...	1072	0.0
gi 10286121 /dataset=dbest /taxon=96...	718	0.0
gi 9872634 /dataset=dbest /taxon=960...	680	0.0
gi 2656674 /dataset=dbest /taxon=9606 ...	549	e-154
gi 9882497 /dataset=dbest /taxon=960...	541	e-151
gi 689641 /dataset=dbest /taxon=9606 /...	525	e-147

## **EXPRESSION INFORMATION FOR MODULATORY USE:**

### library source:

#### Expression information from BLAST dbEST hits:

gi|10934204 Whole embryo (mainly head)  
gi|10286121 Hepatocellular carcinoma  
gi|9872634 Non-cancerous liver  
gi|2656674 Fetal liver spleen  
gi|9882497 Non cancerous liver  
gi|689641 Liver

#### Expression information from PCR-based tissue screening panels:

Mixed tissue (Brain, Heart, Kidney, Lung, Spleen, Testis, Leukocyte)

```

1 MDPMELRNVN IEPDDESSG ESAPDSYIRI GNSEKAAMSS QFANEDTESQ
51 KFLTNGFLGK KKLADYADEH HPGTTSFGMS SFNLSNAIMG SGILGLSYAM
101 AYTGVILFII MLLAVAILSL YSVHLLLKTA KEGGS LIYEK LGEKAFGWPG
151 KIGAFVSITM QNIGAMSSYL FIIKYELPEV IRAFMGLEEN TGEWYLNNGY
201 LIIFVSVGII LPLSLLKNLG YLGYTSGFSL TCMVFFVSVV IYKKFQIPCP
251 LPVLDHSVGN LSFNNTLPMH VVMLPNNSSE SDVNFMDYT HRNPAGLDEN
301 QAKGSLHDSG VEYEAHSDDK CEPKYFVENS RTAYAIPILV FAFVCHPEVL
351 PIYSELKDRS RRKMOTVSNI SITGMLVMYL LAALFGYLT F YGEVEDELLH
401 AYSKVYTLDI PLLMVR L AVL VAVTQTVPIV LFPIRTSVIT LLFPKRPFWS
451 IRHFLIAAVL IALNNVLVIL VPTIKYIFGF IGASSATMLI FILPAVFYLK
501 LVKKETFRSP QKVGALIFLV VGIFFMIGSM ALIIIDWIYD PPNSKHH

```

# **FEATURES:**

## **Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

Number of matches: 5

```

1      83-86 NLSN
2      260-263 NLSF
3      264-267 NNTL
4      276-279 NNSE
5      369-372 NISI

```

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
cAMP- and cGMP-dependent protein kinase phosphorylation site

503-506 KKET

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 7

```

1      33-35 SEK
2      49-51 SQK
3      129-131 TAK
4      290-292 THR
5      360-362 SRR
6      473-475 TIK
7      506-508 TFR

```

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 5

```

1      18-21 SSGE
2      22-25 SAPD
3      129-132 TAKE
4      305-308 SLHD
5      309-312 SGVE

```

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 6

1	95-100	GLSYAM
2	153-158	GAFVSI
3	164-169	GAMSSY
4	186-191	GLEENT
5	296-301	GLDENQ
6	482-487	GASSAT

[6] PDOC00009 PS00009 AMIDATION  
Amidation site

58-61 LGKK

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	79	99	1.125	Certain
2	102	122	2.503	Certain
3	153	173	1.197	Certain
4	197	217	1.785	Certain
5	222	242	2.123	Certain
6	332	352	1.240	Certain
7	370	390	2.166	Certain
8	414	434	1.301	Certain
9	453	473	1.520	Certain
10	476	496	2.166	Certain
11	515	535	2.628	Certain

**BLAST Alignment to Top Hit:**

```
>CRA|145000039337444 /altid=gi|12017941
      /def=gb|AAG45335.1|AF295535_1 (AF295535) amino acid
      transport system A3 [Rattus norvegicus] /org=Rattus
      norvegicus /taxon=10116 /dataset=nraa /length=547
      Length = 547

Score = 975 bits (2492), Expect = 0.0
Identities = 478/547 (87%), Positives = 508/547 (92%)

Query: 1  MDPMELRNVNIEPDDESSSGESAPDSYIRIGNSEKAAMSSQFANEDTESQKFLTNGFLGK 60
      MDP+ELR+VNIEP ++S S +S Y +GNSEK AM SQFANED ESQKFLTNGFLGK
Sbjct: 1  MDPIELRSVNIEPYEDSCSVDSIQSCYTGMGNSEKGAMDSQFANEDAESQKFLTNGFLGK 60

Query: 61  KKLADYADEHHPGTTSTFGMSSFNLNAIMGSGILGLSYAMAYTGVLFIIMLLAVAILSL 120
      K L DYADEHHPGTTSTFGMSSFNLNAIMGSGILGLSYAMA TG++LF+IMLL VAILSL
Sbjct: 61  KTLTDYADEHHPGTTSTFGMSSFNLNAIMGSGILGLSYAMANTGIVLFVIMLLTVAILSL 120

Query: 121 YSVHLLKTAKEGGSLIYEKLGKAFGWPGKIGAFVSITMQNIGAMSSYLFIKYEPEV 180
      YSVHLLKTAKEGGSLIYEKLGKAFGWPGKIGAF+SITMQNIGAMSSYLFIKYEPEV
Sbjct: 121 YSVHLLKTAKEGGSLIYEKLGKAFGWPGKIGAFISITMQNIGAMSSYLFIKYEPEV 180

Query: 181 IRAFMGLEENTGEWYLNNGNYLIIFVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV 240
      IR FMGLEENTGEWYLNNGNYL++FVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV
Sbjct: 181 IRVFMGLEENTGEWYLNNGNYLVLFVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFVSVV 240

Query: 241 IYKKFQIPCPPLPVLHDHSGNLSFNNTLPMHVVMLPNNSESSDVNFMDYTHRNPAGLDEN 300
      IYKKFQIPCPPLPVLHD+ GNL+FNNTLPMHV+MLPNNSES+ +NFM+DYTHR+P GLDE
Sbjct: 241 IYKKFQIPCPPLPVLHDHNGNLTFNNTLPMHVIMLPNNSESTGMNFMVDYTHRDPEGLDEK 300

Query: 301 QAKGSLHDSGVEYEAHSDDKCEPKYFVFNSRTAYAIPIILVFAFVCHPEVLPIYSELKDRS 360
      A G LH SGVEYEAHS DKC+PKYFVFNSRTAYAIPIIL FAFVCHPEVLPIYSELKDRS
Sbjct: 301 PAAGPLHGSGVEYEAHSGDKCQPKYFVFNSRTAYAIPIILAFVCHPEVLPIYSELKDRS 360

Query: 361 RRKMQTVSNISITGMLVMYLLAALFGYLTFFYGEVEDELLHAYSKVYTLDIPLLMVRLAVL 420
      RRKMQTVSNISITGMLVMYLLAALFGYL+FYGEVEDELLHAYSKVYT D LLMVRLAVL
Sbjct: 361 RRKMQTVSNISITGMLVMYLLAALFGYLSFYGEVEDELLHAYSKVYTFDTALLMVRLAVL 420

Query: 421 VAVTQTVPIVLFPIRTSVITLLFPKRPFSWIRHFLIAAVLIALNNVLVILVPTIKYIFGF 480
      VAVT TVPIVLFPIRTSVITLLFP+RPFSW++HF IAA++IALNNVLVILVPTIKYIFGF
Sbjct: 421 VAVTLTVPIVLFPIRTSVITLLFPRRPFSWKHFGIAAIIIALNNVLVILVPTIKYIFGF 480

Query: 481 IGASSATMLIFILPAFVYKLKVKETFRSPQKVGALIFLVVGIFFMIGSMALIIIDWIYD 540
      IGASSATMLIFILPA FYLKLVKKE RSPQK+GAL+FLV GI FM+GSMALIIIDWIY+
Sbjct: 481 IGASSATMLIFILPAAFYKLKVKKEPLRSPQKIGALVFLVTGIIFMMGSMALIIIDWIYN 540

Query: 541 PPNSKHH 547
      PPN HH
Sbjct: 541 PPNPDHH 547

>CRA|114000033649823 /altid=gi|10945621
      /def=gb|AAG24618.1|AF298897_1 (AF298897) amino acid
      transporter system A [Homo sapiens] /org=Homo sapiens
      /taxon=9606 /dataset=nraa /length=506
      Length = 506

Score = 597 bits (1522), Expect = e-169
Identities = 315/549 (57%), Positives = 383/549 (69%), Gaps = 46/549 (8%)

Query: 1  MDPMELRNVNIEPDDESSSGESAPD---SYIRIGNSEKAAMSSQFANEDTESQKFLTNGF 57
      M E+ +I PD++SSS S D SY +++AA+ S +A+ D E+Q FL
Sbjct: 1  MKKAEMGRFSISPDEDSSSYSSNSDFNYSY----PTKQAALKSHYADVDPENQNFLLESN 56
```

FIGURE 2, page 3 of 4

Query: 58 LGKKKLADYADEHHPGTTTSFGMSSFNLSNAIMGSGILGLSYAMAYTGVLFIIMLLAVAI 117  
 LGKKK Y E HPGTTTSFGMS FNLSNAI+GSGILGLSYAMA TG+ LFII+L V+I  
 Sbjct: 57 LGKKK---YETEFHPGTTTSFGMSVFNLSNAIVGSGILGLSYAMANTGIALFIILLTFVSI 113

Query: 118 LSLYSVHLLLKTAKEGGSLIYEKLGKAFGWPGKIGAFVSITMQNIGAMSSYLFIIKYEL 177  
 SLYSVHLLLKTA EGGSL+YE+LG KAFG GK+ A SITMQNIGAMSSYLFII+KYEL  
 Sbjct: 114 FSLYSVHLLLKTANEGGSLLYEQLGKAFGLVGKLAASGSITMQNIGAMSSYLFIVKYEL 173

Query: 178 PEVIRAFMGLEENTGEWYLNNGNYLIIFVSVGIILPLSLLKNLGYLGYTSGFSLTCMVFFV 237  
 P VI+A +E+ TG WYLNNGNYL++ VS+ +ILPLSL +NLGYLGYTSG SL CMVFF+  
 Sbjct: 174 PLVIQALTNIEDKTGLWYLNNGNYLVLLVSLVILPLSLFRNLGYLGYTSGLSLLCMVFFL 233

Query: 238 SVVIYKKFQIPCPPLVDHSGVNLSEFNNTLPMHVVMLPNNSESDVNFEMMDYTHRNPAGL 297  
 VVI KKFQ+PCP+ + N + N TL ++P  
 Sbjct: 234 IVVICKKFQVPCPVEAA--LIINETINTTLTQPTALVP----- 269

Query: 298 DENQAKGSLHDSGVEYEAHSDDKCEPKYFVFNSTRAYAIPILVFAFVCHPEVLPIYSELK 357  
 + + +D C P YF+FNS+T YA+PIL+F+FVCHP VLPIY ELK  
 Sbjct: 270 -----ALSHNVTENDSCRPHYFIFNSQTVYAVPILIFSEVCHPAVLPIYEEELK 317

Query: 358 DRSRRKMQTVSNISITGMLVMYLLAALFGYLTIFYGEVEDELLHAYSKVYTLTDIPLLMVRL 417  
 DRSRR+M VS IS M +MYLLAALFGYLTIFY VE ELLH YS + DI LL+VRL  
 Sbjct: 318 DRSRRRMMNVSKISFFAMFLMYLLAALFGYLTIFYEHVESELLHTYSSILGTDILLIVRL 377

Query: 418 AVLVAVTQTVPFIVLPFIRTSVITLLFPKRPFSWIRHFLIAAVLIALNNVLVILVPTIKYI 477  
 AVL+AVT TVP+V+FPFIR+SV LL + FSW RH LI ++A N+LVI VPTI+ I  
 Sbjct: 378 AVLMAVTLTVPVVFIPRISVTHLLCASKDFSWWRHSLITVSILAFNTLLVIFVPTIRDI 437

Query: 478 FGFIGASSATMLIFILPAVFYKLKLVKKEFTRSPQKVGALIFLVVGIFFMIGSMALI IIDW 537  
 FGFIGAS+A+MLIFILP+ FY+KLVKKE +S QK+GAL FL+ G+ M GSMALI++DW  
 Sbjct: 438 FGFIGASAASMLIFILPSAFYIKLVKKEPMKSVQKIGALFFLLSGVLVMTGSMALIVLDW 497

Query: 538 IYDPPNSKH 546  
 +++ P H  
 Sbjct: 498 VHNAPGGGH 506

**Hammer search results (Pfam):**

Model	Description	Score	E-value	N
PF01490	Transmembrane amino acid transporter protein	187.0	2.9e-52	2
CE00398	E00398 CD53	4.0	4.8	1

**Parsed for domains:**

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00398	1/1	90	110 ..	1	23 [.	4.0	4.8
PF01490	1/2	99	236 ..	1	179 [.	58.9	2.5e-14
PF01490	2/2	305	529 ..	200	467 .]	133.9	3e-36

1 AGCTTAGCAA TATGGATCAA GAGGTCCAAT ACCTGATTAA TAAAAGTTTC  
51 AGGAGTAAAC AAAGGGGAAG AAATAGTTTT TTTAAATAGT AGAACTTTTT  
101 TTATTTTTAG AAAATGTGTC TTCTATAGAA GAAAGACAAG CCTTTTGATT  
151 GGGCCGTCTG CATGCTGAGT ATGATGAATT TTTAAAGCGA CTCACATCTA  
201 GTCACGTCGT GATGAAAGGA TAAGGATAAA AATTCTGAAA TCCTCAGAAA  
251 ACCATCGATA AATTATCTAT AAAGAAATAA GAGCCAGACT CATCAATAGA  
301 AGCTAGAAGA GAGAAGTTTC TTCAATATTC TGAAGGAAAA TGCTTCTGAA  
351 TCTAGAATTC AAACAATTAA CAAAGTTTGA AGGCAAAATA AAGAATTTTC  
401 CAACATGAAG CAACTCAGAA ATTCTATTTA CAGACATAGG CTCATTGTGT  
451 GAAAAAGTT ATTCAAGGCA TTATTTTAGC ATAATGCAAA ATAACTGAA  
501 GAAAGAAGAT AGAATGCCGT TCAAGAACT AGCAGCTGAG CAAGACTCAG  
551 AGGTTGGAGG AGGAAGCCAT TCAGAATGAG AAAGAGCATA GAAAATTTGC  
601 TTTCAAAGTT TTGGTAATAT AGAATTATAT TTCACTTATT ATGTAGTCAA  
651 ATACACCACT TTGTCTTTAG GGCATACAT TTATACAGTG ATAATACTGT  
701 AATTGCTGCT TATTGGTTTT CCATGTTTTAG AAACAACCTA CAGGCAAGTT  
751 ATGACACTTG TTTCACAGAA CAAGATGAAA ATATTATGAT TCTCAAATTG  
801 TAAAAGTATT TTATTAACCTA AAATAATTAG GAGTGTAGGA GAAGGAAGGA  
851 AAGAAAGAAA AAGTATGCTA ATGTCCTTAT TTTTATGGG TAACCAGTCT  
901 AAAATCAGTA AACCAAGTCA AAAAAGCTTT AGTGAATTAT TCAGATCTAG  
951 AATGGCTAAC TTTAAGTAAC AAGCTAAAAA CAGAAACCGT CAATAGTGGT  
1001 TGCTGCTGGG AAGTGAGACT GGTACTGTGT GAAGAATGAG GAAAACCTTT  
1051 GTACTCATTT AGTGAGTTTC TTTTTTTTTT CTTTTACCCA TATGCATGTC  
1101 TTAATCTAT TCTCTCTTAG CTTTTAACCT GCTTCTTTTC ATCTTTTATG  
1151 TATATACATT TAGGCTGCCT TATATTAATA ATAGTTTCAT TTTTGTCTCT  
1201 CCTGCTTAAA AACTGTGTG CTATTTTTTT AAATTCTGAG AACTGCTTTC  
1251 TTTATTTCTA GACAATTCTC TGCCATTATC TCTTCTGTT TTGTCTCACC  
1301 CTAGTCTCAC AATTCTCTAT ATTGGAATGA CTATCAGTGT ATATTTGAAC  
1351 TTGTAATTCT TATTTTTTCC CCATTCCTCT TAACTTCTTA TTTGTATTTT  
1401 TCTTTTTTTT ATCTCTTCAT GCTATAATTT GAGTGATTTC CACAGATCTG  
1451 TCTTTCAATT TTATAAGTCT TCCTTCAGCT GAGTTTTTTT AAATTTCAT  
1501 GATTCTATTT TTTCTTTTTT TTTAAGAATT CCTTTTTTTG ACTCTTTTTG  
1551 CAACAGCCTG TTCTCCTTTT ATATTCCTTT ATAATGTTTT TATTCTGTGA  
1601 AAGTTATTCT CTTATTTTGA ATGTTTTCTT TCAAAATGTC TTTCTTTTAA  
1651 TTAATTTAAT GTAAAAGTCC CTTTTAAATT GCTTTGTTAT TTGTAGTTCC  
1701 TTAGATGTGA ATTTTATCAT TTCTTGTCCT TACTGGCACT CTTGCTAGTG  
1751 AGTTTCCATG TGTGTTCTAT ATGTTTTGTA ATTTGAGGAT GTGAACTTTT  
1801 CTCAAGTGTG AGTTGCCTTT CAAAAAAGTA CTGCCATGGC ACTGGGTTGT  
1851 GGAGGTATTC CCATGTGGTA GTTCTGTTT GTCAGAGGAA TAGCACATTT  
1901 TGTGACTTCT GGAGCAATTT TTATGTTAGT TTCTCTGCTC AAGATTTCCT  
1951 TATCAAATGG GTATTGCACA TGTCATGACC AACTTTTTCA AGAATGATAG  
2001 TGTTTCTCCT AATACGATGG TTCAACAATA ATTGAATGAA TCTAATGGTA  
2051 AGAATTTTCA AAGAAATTAT ATCAACTACA TATAGTAGAT TCAAGGCATT  
2101 TTTCAAAAC ACAATGCCAG TCCACCCCTT TTCACTATAC AATTGAGGAA  
2151 AATGAGGTCC CCAATGTGA AATGACTTCT GCTGAGATCC AATGAATTAA  
2201 AGGCAGAGCA GAGGCTAAAA TCTAGATCTC TTTGTTGTTA AAATACATTT  
2251 TAATTTGACA CAGATGATGA GTAAATGCTGA CCCAGAGGTA AATCTGAACT  
2301 TTCTTTTGTT ACTATTCTTA ACTTTGGCTT CAGGATCCAA GTGCCTAGAA  
2351 AGTTACTTCC TAAACTTGAT CCTCACCTAT GTTGCATATT ATCAAGCATT  
2401 TGGTGGTGTT AATCTTTTCA TGTCCAATTA AATTAAAGCA GTAATTTTCT  
2451 TTCTAGTTAT TGCTAGTAGA GACACTGGTA GATTCTGCCT TGGTAGACCT  
2501 TCCTCTGTCA ACAATTTACT TTTGTCTTCC TTTCTTTTAA AACATGTATC  
2551 CCACTCACAA ATACCTAAAT TTCCTTGAAG ACTGCTGCCA TGTTTTAAGA  
2601 TTTCTTTTTT TTTCCATAGT GACTAGTAAA ACCTGCCATT TTCATTATAC  
2651 ATAGGCATCT TATAAATATC TGCTAATTTA GCAATTATTA GTAATTTCTT  
2701 TTCTTCTCTT CCATTTCTTC CTTTCTTGTA TTGGGTAAAG GAACATTTCA  
2751 GGATTTTGCTT ATGTAAGTTT TTCAGGAGTT TCTTTCCTTC CTCCTTTTAA  
2801 CAGAGAGCAT ACAAATGTGA GATGATTCAT ATTCACTTAT TTCATTTAAA  
2851 TAAAATTATA ATGATGTATG TTGTGTTCTG TTTGCAGAAC AGAGTGTCTT  
2901 GAACATCAAC ACAAAGTGA AGAACCTTAA GCTGAAGGTA CAGTATATTA  
2951 TTTACACTGA AGGGGCTTGT GTGTGGACAA GAAAGCGCTG ACAGCTCAAA  
3001 TGGATCCCAT GGAAGTGAAG AATGTCAACA TCGAACCAGA TGATGAGAGC  
3051 AGCAGTGGAG AAAGTGCTCC AGATAGCTAC ATCGGGATAG GAAATTCAGA  
3101 AAAGGCAGCA ATGAGCAGGT ATGGGGTTAA AAATTACTAT GTTCCATGGA

FIGURE 3, page 1 of 23

3151 AAAATAAGAC AGGATGTGGA CATGGAAAAC AGGGTCTTGA TGGGAAGAAC  
3201 TGGATTTATT ACAGGTAAT TTGTGATAAC AATGATATTG ATGCTAGCAC  
3251 ATCAATTCCT TGGTCTGAA ATACAGTGAT AATGTCAATC TCTTTTGTGA  
3301 CTGATTTAGA ATTGAGGTTA CAATGTCTTT GTCTCCATTA ATAATGTGTA  
3351 ATAATTTTAA TTATTTTAGC CTATTGCTCC TCTTATCTTT CTCAGATTCC  
3401 TCTTTGAATG TTGCTACACC TCCTGGTTTC TGTAGGGATT CTTTCTCTC  
3451 TAAAAGTATC CTCTGGGCAA GCTCACTCAC AACTACTATG GCCTCACCCCT  
3501 CCAAATATAT GCCATATACC CAGCCTGTTA AGTTTCTCTA CTGAATTTCA  
3551 GATAATTATA TCTGAATGTC TACTGCACGT CTCTACTGGA CCATTACTGT  
3601 GTCTAAATTG CCTCATTAT AAAGTTAAAC CTGTAATGTC TAATACTGAA  
3651 CTCTATCTT TCCCTCCAAA ACCTGCTCCT CCTCTAGTAA TCCCCATCCT  
3701 AGTGAAATC ACTGCTATCA TGTAGCAACT CACTCAAAAG CCCCTAGGTG  
3751 TAAACTTTGA CCCACATAGC CAACGGTCAG TCATATCCAG TTGGTTTGAC  
3801 CTTATTAATG CTTCAAATAC ACCTACTTTT CTGTACCCAT TCTACTGTGG  
3851 TCTTACGTTA GGCCTACATT AAATGTGAGA CAGGGAGAGA GCCCTGATTT  
3901 CTCTCCCTGT CTTACATTTT GCTCTCCTCT GTCTAGCCCT CTACACTCCT  
3951 GCAAGAGCAA TCTCTTACAA TTGCAAATG AATCAATTC CATCCTTAGA  
4001 TAAAGCCCTT CTGCACCTCT CCAATAGCCA TAAGAGAAAG TAGATTACAC  
4051 AACTGCTGG GCACGTAAGG TCCTTTGTGA TCTGTCTTG ACCTGCCCCCT  
4101 CCGTGCCTGT TTTTGGCCCT CTCCCTATTT GTTACTTGTT GCCTCACTC  
4151 ATCTGCTCC AACTGCCTGG AATCAGTCAC CTGCTCCCCC TTTCTCCGTG  
4201 TTGACACCTC TCATCCTTCA AGAATCAGCT CAACATCAGG TCTCCTATGC  
4251 AGCCTTTTCC AAATTACTCT ACTCCCCAT GTAGAAGTGA CTGCCCTCC  
4301 TTCATGTACC CTCTCCCTGT GCAGATGTTA ATTACGCCAC TACTACAGGT  
4351 TAATGGCCTC TGTGGTCCCA CCACCTGCCA CATTGTCTGG TGCATAGTGA  
4401 GTGCACAATA GTTATTTGAT AAGTCAATTG ATTTCCACA AAATGTTATA  
4451 TCAAATTGTA CATGATTTAA GATGCTCAGA AGGGAATTTT TGACCAAATC  
4501 TAGGCGTGAA ATAGAGAATA TTGTGCTCAA ACAAAGACTT CTCATTTTAT  
4551 TTACAACACC CAGGAAAATC CATCAGGAGA AACTACCGTT CTTCTTCAA  
4601 GTAGTCAAGT GCAATGAAC TTAGGGATGT CGGACTAGAG AGGCCACTGA  
4651 GATGTAAATT ATAGCATTTT CTAAATTAGG TGACCCTTGA AGAAACACTA  
4701 GGGTGCTAGA AGACAGGGCT TTGGAGTCTG CAGAGTAGTT GCCTGACTTT  
4751 AGAGAAGCTG TTTGTCTCTT TTGAGCTTCA ATGGAAAATG TAAAATGGCA  
4801 AACCACAGC TGCTTTTCAA GGATGAGATG GGTGACCAGA ATATAGATGA  
4851 CATTCATAC TTTTATTATA CTCTCCTTC ACTGCATTAC CCTCAGTAAA  
4901 TTGATTCAAA CTTGAGGATG TTTCTGAAAG GCATGCACAC AAATATGAGC  
4951 TCTGCCGAGG TTGACAGAGT TAAAGGGGAC ACCCTCCTAA GAACTGTCAT  
5001 AGTGCATTC CACTTGATCC TCAAAAGCCA GAGTAGAAAG AGCATGAATG  
5051 CTTTTCTTAA GCTTCATGCA ATGTGTTCCG AACCCTCAC AGTGACTTAC  
5101 CTTTTATCTC CTGGCTTAAA CATAGGACAT CATTTTGCAG TTTTAAAAAT  
5151 CAGTTTAAAG AGATGGGTTT TATCTATGTG TGGTTTGGAT TGAACCCCTA  
5201 AATGTAAAT TTTGAGAAAT TCAACATAAT GTATTTATTT GTGATCATTA  
5251 TACTTGTGTT TTCAATACAT GCTGGGTTTG GTATCAAAAC ATTTAACATA  
5301 CTGGGGACAT TTCTCATCTA TTTTATACAA TCTTGGCATG TTAAATGACT  
5351 ACAACTCATC TCATGCCAAA ATAAGAACAT GCAAATGCCT CAAAGAAAGA  
5401 AAATCTGTTT ACTTTCAAAT TCTCAATTTT AAAAACTACT ATGGAATACA  
5451 GATTTTAGTT TATTGATTAA AATAAAGATT CCAGAGTTTA AATTCTAGGT  
5501 GGCACTTTGG TTTTATAGT CCTCAGGCC ATTTTAGGCT TCATTTTATC  
5551 CTGTCACTC AGTCTCCAAC TGTGAACATT ATGTACCAGT CTTACATAG  
5601 CAGGTACATT AATTACAGAC CATTAATGTA AACCACAAA GAGTGGTGGG  
5651 CAGTGGGTGG GGGGTGAATG GAAATGGAAA GAGGCAACAA CTGAGGGCAT  
5701 TGTGCTTTCT GTGAGAAATA TGGGGAGAAG GCTAGGAAAT GTTCTTAACT  
5751 TGTGTACTCA GAGCTATTTA TGCCTTGAGT TCTAGAAAAG CACATACAC  
5801 TTTGTGGTTT CGTGTGCTGT TTCTATCTAC ATCTCATACT GTTTTCTATT  
5851 CTCAAAAAGT AACCTGTCA TCCTCTTCC TCTCCAGATT ATTTTCAGGA  
5901 TTAGCTTCTG TTATAAAAA TAGCTTGTA AGATCTCCTA CAATAATTAT  
5951 TTTCTATTTT ATTTCTAAGG TTTATTTATT TATTTATTGA GACAGACAGA  
6001 GTTTCACCTC TGTGGCCCAT GCTGGAGTGC AATGGTGCAA TCTCGGCTCA  
6051 CTGCAACCTC TGCCTCCAG GTTCAAGCGA TTCTCTGCT TCAGCCTCCT  
6101 GAGTAGCTGG GATTACAGGC GCCTGCCACC AACTCGGCT AACTTTTTGT  
6151 ATTTCTAGTA GAGACGAAGT TTCACCATGT TGGCCAGGCT GGTCTTGAAC  
6201 TCCTGACCTC AAGTTATCCA CCCACCTCAG CCTCCCAAAG TGCTGGGATT  
6251 ACAGGCGTGA GCCACTGTGC CTGGCTCTA GGATTATATT AATAGAACAA

FIGURE 3, page 2 of 23



```

6301 TCTTCAATTA TTTTATCTTT CTTTATCTTT CTTTTCATGT AGGAAATGTC
6351 CTAATAATTT CAAACCCTCA ATTTGAAAGC ACTTTTAAAA TCATACATAG
6401 TCGAGCATTT TATATAAAAA CAACTAAAAA GTCTGTGACA TTTTGCAGTA
6451 TAAAAATGCA ATGGCAGCAG CAGGCCTTAT TAATTGAGCC TCTTGGAAAT
6501 GTGGCTGGTC CTAGGTCCGT AGCCTCAAAG GCCCTGGCTT GTAAGTGCAG
6551 GAGCTGACCA GCACAGCTCT ATAACCAAGT TGTACATCTT CTAGCCTGTG
6601 TCCAAGAAAA CCAGAATCAC AACGCTCTGT GGATAGTGAC ATCTTAAAGT
6651 TTTCTTTCCC TCCCAACTCT TTTGCCAGTT CATTGAATTG CTTTAATAAT
6701 TTCCTTAGTT TCATTCAATTA TCTGTTAATA ATCCATGTAC ATTTTGAGAG
6751 TAATTAATAA ACATACGCAC ACACAGAAAC AACCAACACA ACACACAGCT
6801 ACCACTGAAT TACTTTCCAG TAAGAGATGT ATGTATAAAT GATTGTACCA
6851 AAAAAAAAAA AAGAAAGAAA ATACCAGCTA CAGGGCCCTG CCTGGGACTG
6901 CTTGATGCCA GGGGGAGAAT GGGGTCTCCC CCTGGGTATG GGTGGGTATG
6951 GGCCTGCTGC TTCACCTTTC TGAGCCACAG TTCCCTATAG GGATATTTTG
7001 AACATCAGAT GAGATAAGGA TCACAGTGCC TAGGCATTTA ATAAATATTC
7051 GTTGAATTA TAAATCATC TGTATTGGT ATGGTAGTAG TTCAGAAAAA
7101 TCTGTCATAA CCCTGTACTC TTTCTTTGGA AGGGCTCTAA ATGGGAACAC
7151 AATTAGTTGT AGTCTCTTGC ATAGCTAATG TGAGAAAGAG GGAATGTGGT
7201 ATAAACAATT TTTTAACTAA AAATAATATT TCCTTCCTTT ATAACATCCT
7251 TCTTCCATCC CAAAGTATAG TTGTAAATGG AACTCAAAAT TGTTGGTCTG
7301 GAATGACCGT TAGTGTGAAG GAGGAAAAGA AAATTGGGGT GTCTTATTTT
7351 CCCTCCTCTG ATTCAGTTAC TTAGATCACC TGAAACATAC ATATGATTCA
7401 GAGCATATAT TTAGATGTTT TCACCTTCTT ATTTGTGTGT GTGTGTGTTC
7451 AGTCAATTTG CTAATGAAGA CACTGAAAGT CAGAAATTCC TGACAAATGG
7501 ATTTTGGGG AAAAAGAAGC TGGCAGATTA TGCTGATGAA CACGTAAGTG
7551 AATCTATGCT TTCAGGCAAT AAACGGGACT GAGGGTGTCT GATCTACCTA
7601 GGTCTCTGTG GGAAACAAT GTGACTGAAA TTTTCCAAGC CTTGATCAGC
7651 ACATTCTGTG TTTATTCCAG CTCTTACTGG AATAAGGGCT TGTTTTTTCC
7701 TGTTCGCCAT ATGGCTGCAT GAATCATTTA TGAAACTTAT GTGTTTTGGG
7751 GGGAAATCAT TCTAACCCAA AGGTAATCTA CAATCATACA TGTTTTCCCT
7801 TCTTTATGTG ACTCCCCTTG TAATTTGTAT TTTTACTGAG GCCTCTGCTG
7851 AAACCAAGCA CTGCATTCCG TTGAAAATTA CATGCTTTTA TTGATGTTGA
7901 GTAATGGCTT TACTCCTGTA ATGTTATCTT AGTCTCAAT TTTGGACTGT
7951 AATCTGCAGA TAATGTGAGA ATAAGGATAA CCCCTAAAGG TATGCCCTTT
8001 GGCAATGTT TGCTTATAAT ACATCCCTTC TTTTCAAGC ATCCCGGAAC
8051 CACTTCCTTT GGAATGCTT CATTTAACCT GAGTAATGCC ATCATGGGCA
8101 GTGGGATCCT GGGCTTGTCC TATGCCATGG CCAACACAGG GATCATACTT
8151 TTTATGTAAG TGAATGTATA TGTCTACATT TGGTGATGAA GTCCATGCAT
8201 ACCTGGTGGC TTTTCAATT AACAATCTCA AGTTTGATCT TTGTGAACGT
8251 GAAGACTCAG AGGAGGCTAA TCATGGCACT TGGTCACCCA ACCATCCCTA
8301 ACCCAACGGC AGAAAGTGTA TGTGCTCAAT CAACCAAAGT GCTGGAGCAG
8351 CCTCGCCAGA AGAATTTTGT TATTCAGTAA ATACTTGAAA TAATTGGTG
8401 TTTAGCAACC AAAAAGATCT TCCCAGAAG CAAATCTGAT TTTATCTCAT
8451 TCTTAGGAAA GAAGCAACCA AGCCTAAGAG CCCTGCATGC CCTTGCCTAC
8501 CTTATGTCCC ATTCCCTGTA CCCCTGTGCG ACAGATACAC TGGGCACAAT
8551 AGCCTTCTCT CCATCCTATG AAGATGCCAC ATTCCCTCTC ACCATGGAC
8601 CTTTGCACAT GGTCTTGGA CCCTCTTCTC TTCCTTCTTC ATCTAGTTAA
8651 CTCCTCATAT GTCAGTTCAG TCTCACCTGA ATACTGCGCG CCCTGATCTC
8701 CATGACTGGG GCAAAATCACC TTATCATAAC ACTCACCACA ATTTTAATGT
8751 TTTAGTGCCA TTTGTCTGAT TCATTTGGTT AATATCTGTC CCTCTTGCTG
8801 GACTATAAGC TCTAGAAAGT TGAGCCCATG TCTGTTTTTA CTCACCAATG
8851 TCTCTACCTC CAAACCTAGA GCAGTGCTG GTACAGGCAA TATTTGTGTA
8901 GTGACCAAAC CTTATTCTTA AACCTACGTA CTTTCACCAA ACTTGTTCAC
8951 ATGCTGCCTA AGGGTAGCAG CATCTGGTAG TTGACCTGTA GGGTGGATAC
9001 TGCATGTCT ATGACAGACA ACAACAGACG TTTATGTGCA TCATGTACAG
9051 CCTGGCATT TCCAGGATAT AGTTGGCAGC AGTGGAATTC TTCACAAGAA
9101 TAAAGTCTGA TGTTAGGCAC CACTGTGGAC ACAGATCCTA ATCCCAAATG
9151 CAACGCTAGA GAGTTAAATA ACTGTCTAAG AATGCAACAT TTATATCACA
9201 AATATGTGCT GTTTATGTTT TGAATATCAC ATATGATTAG TAATCACACA
9251 GCTATTTGAG GGCTAAGCAT CAGGACTATA AATATTTGTA TTGTGTTAGT
9301 GCTTTGATTG AACTCTTTTA TGTATAATAT TCTTCAGCTG AATGGGTTTT
9351 TATATCAACT TTACTTTTAT ATAAGCCATG TTTTGAAATA AACTAGGATT
9401 TTAATAATCT GAATTTTAAT AGCTATGTAT GTAGTCATAT ATTTGTATGC

```

FIGURE 3, page 3 of 23

9451 TTTTGTAAATG TGCTTACCTC TAAGACAAAA AAACCTGCCT TTCCTTATTA  
9501 ATTATACATA CCATTAAAAAT GAATTAGGAA GTTACAGATC ACTGATGAAT  
9551 AGAAATAGGA AAAACTTCCC CCAATCCCAC AGTCATAGAT CATCTTCATG  
9601 AGAGAAGAAT GTTCCACTTT TTAATAATGAG GGCTTCATTT TAGGCTTATA  
9651 AACACTTAGC AGATGAATTT GGTGAGAACA ATTAAATCAC TAAACATCAT  
9701 GGGGTGTGTT TTGTGTGTCT AAGTAGCCCA GACTGGATTA AGCTTTCTCT  
9751 CTTAATTTAT AGCAAGTGAC ACAGTATTTT AAAGGTTTTA CTCTTAGTAT  
9801 TTTCTGCCAG AGAAAGTACA TGTTTAGAAT ACAGGGAATG CTCATTATTT  
9851 TTCCAGGGAA CAAAATTATA TAATCTGAAT TACATTATTC CTAAAAACA  
9901 GTTAAGTTCA TAAGGCATAT GGAAAAATAT AGGAATAAGT CATTGGTTAG  
9951 ACAGTTCTGG CAAACATACT CTATGGAAAA TAAGAGTGCA ACATAGCTAC  
10001 AGGGGTTATA AAATTTATAA TTCATGGTCC AAATGTACAT TTGTAGTATT  
10051 GATTTTCATTG GGAATTACCA AGGGATTAGA TCAATTGTGG GGAAAGTGTA  
10101 TTTTTTAAAA ATAAACAAAG ATAAAGATTT TTTTCTGAA TTCCAGGTAA  
10151 AAGGCAGCAT TGCTCCTCCA TTTATTACGT AGATGCTTCT ATCAACATTC  
10201 TTATTTTGT GCTCCAAATC TTGGATTGG AAAAATACCA ATCCGTATAA  
10251 ACATAAGAA ACCATACATG CATGTGGGGA TCCTAACACC AGAAATGACT  
10301 CTGAATGCAA AAAAAAAAAA AAAAAAAAAA GGGAAATTTT GTGCCCCATC  
10351 CTTAGCTTTC TCTGCTTTCT CTATTATATA TGCAACTGCC TGCCCCCTCTA  
10401 TCTTACAAAG TACTTCGTAA TCTAATGCAC AGGATCAGCA GTAATGCAGC  
10451 TCAGACTGCA TGCTTTCGCC TTTGGATTCC TAGATTTTCT ATTAAGGTTT  
10501 AGTCAGGCTA TTGAATAGCC CTTCAATTCT AAGTGCTGAT GTGAATATCA  
10551 TGCAAAATATG ATGTACATAT TCCCATGTGC TGAGTAAGTA GATGTAGCAT  
10601 TTGCTAATGT TGCTATACAT TTAGCATCTA AGTTATGAAC CAGATTCTAC  
10651 CACTGGGTAA CATTAAAAA AAGTTAGGGA CTTCAGGTAT GTAAAATATA  
10701 GCAAATCTTA TTTCTACGAC TTTAAAGGGT ATGTGTAGAG TTCTGAAAAG  
10751 AATTTCTCAG CCTCCCCCAA ATCCACATAC TTTTGAAAG CTGATGATTG  
10801 AAAAGATTAA TGTGATCCTT TATTGTAACA TCTAACATAA TTACATTTTA  
10851 TTTATTGTAG AAACCTTTAT ACCTACTCTC TCTTCCCTTT GCAGAATCAT  
10901 GCTGCTTGCT GTGGCAATAT TATCACTGTA TTCAGTTCAC CTTTTATTAA  
10951 AAACAGCCAA GGAAGGAGGT ATGCTACCAC TTGAGTCCAA CACATTCTAT  
11001 TTTAATCTCT ATAAAAGAGT ATTTTCAGTCT GTTGCTTCAT AACCTTAGGA  
11051 TGATTATAGT CAGTTTCACA TTTCAATTTT TTCTGAGCCC AGTGACACGA  
11101 TCTCTCAGTG TTTATAGTTG TTTGGGCAAG TGAGAGGCAG GAGTGAAAGT  
11151 CAACTGGCTC AGGTTCAAGA CAAATAGAAA AAAGAAATTT CTGATATATG  
11201 ATAGAAATAA CTGTTTTGAC TTGCTACATG CAGCTAAAAA AAATAAAACC  
11251 ATTGATTCTT GTTTGGAGAA CATTTTGATA TATTGCTTAT TGGTTTTTGA  
11301 GGTTGCATCT TTTGGGCTTA TAATTTCTAT ATGATGTTTA TTTACATGTT  
11351 TGAGACTCCA GCATGGAATT ATATGACAAA AATATTTTAG TCATTAATAA  
11401 AATCTCTTTA ACAAGGCTAT TTTATCTTTG ATTGTAGGGT CTTTGATTTA  
11451 TGAAAAATTA GGAGAAAAGG CATTGTGATG GCCGGGAAAA ATGGAGCTT  
11501 TTGTTTCCAT TACAATGCAG AACATTGGAG GTAAGGGGAT ATACTTTCCA  
11551 ATGGATCCCA TAACTTTCT ATAGCGTGT CAATAAATAA GAAAACCTAT  
11601 GGCAATAAAC AGGCACTTTA GATACAGAAA AATTGCTACT TATAGTTCTT  
11651 AAATTTTAAA ATGATAGTTT CTTAAATAGG TTTGTGTCTT GCTTTAATTA  
11701 AAAACAGCAA TATCTAAGAA TGAAATAACA TATAAAACCC TGCCAATTGA  
11751 ATTCTAGAAT TAAAATATAA AATAAAAGCT TTCTTGATT TTAATGTTAT  
11801 TATAGCATGA ATTATTACTC TTAATAATTT AAGAATTTGT GCTTATATCT  
11851 GTCATTGACA AAACAGTTGA CGTTTTCTAT GTGTGACTGA GTTCGATTTA  
11901 CTAAACTGAA AAGTGGGTGT CTGGGGGAAC ATAGCCAAAT GCTGTGGTCC  
11951 TTGAAACGCA GCCTGCACTG AGCCAGCCCA CTAGACAGTG TCTCTGGAAG  
12001 TTTACTAAG CAAAAGTCTG GCTAGGCATC AAATGCACTA TAAACCCCGG  
12051 TTTGTTGATT CTATGGATT TATAATTCC CACTGAATTA TCATTTCCAG  
12101 TGTAGGACCT AGAAATATAT ATATATATTT TTAACAATGT TCTCTCGTTG  
12151 GTGTGTTTGC CCACCAGCTT CATACTGTTT CTGTTGTGTC TTTGGCCCTC  
12201 AGAAGGCATC CAAACCCATA TTTCAGATGT CCTGCCGGCT GCTTCCTGGC  
12251 ACATGGCCCC AGCCATCTCC CCACATAATG ACACTTACTC CCTCACCTCC  
12301 TACCCAGTCC CTAAACCTGC TATTCTATTT CTCTGATCTT TCTTTTCTCA  
12351 GTGAATACCA CCAGCAGTCA TCCAGTTTCT GAGGCGAGAA ATCTGGATGT  
12401 CAGCGTAAAT GTTTCCTTTT CCCCACCTCT GCATGTCCAA TCAATGGCA  
12451 AAGTCTGTTC ATTTGATCTC TTAATTATCT CTTGAACCTC TCCTCTCTGT  
12501 CCGTCTCAT GACCACAGAT GATCACCATT TATAGCTCAG ACTATTGCAG  
12551 TAGTCTCTA ACTGGTCTTC CTGGCTTGAG TTTCCCTGCT TCTCAGATAA

FIGURE 3, page 4 of 23

```

12601 ACTCTAATTT GTTCTCCAGA TAAACTTTCT CAAATTTGAG TCTGTTTCTA
12651 CTTTTGTGCG GCATAAAATT CTTGAGCATG CCTTTATTAT TTTCAGGAA
12701 AAACCTTAAAC TCATTGGACT GACACAAGAT CTTCGTCTAG TTCTTCTGCT
12751 CAATCTTTCT AAACCTTCCT AGCAATGCC ATATCTATCT ATCTTTATCT
12801 ATCTATCTAT CTATCTATCT ATCTATCTAT CTATCTATCT ATCATCTATC
12851 AATTTATCCA TCATCTATAC CCTACATGTC CTGTGTCAAA CCATAACAAA
12901 TTATATTTAT TCCCCTAACA GTACTATTTT AATATTTTTA AAAATCATCC
12951 ATGCCTTCTT TTCACAGGCT ACTTTCTCCC CTTGACTGTC TCTCAAAGTC
13001 CTCCAACCTT AACACACACG CACACACACA CACACACACA CACACACACA
13051 CACACACAT TTTCTCTCA CTCTGCTCAC CTGGTCTATT GCTCCTCTAG
13101 ACTGGTAAAT ACTAGTTCCT CTGGGCTCTC ATGGTCTGT TTGTATCTAG
13151 TATGTTACTG TTTTCTAAAG GATATTTTAA AACACTTGAG TAGAGAATAA
13201 GCTTTTGGAG TCTGATGGAC CTGAATTTGA GTCTGTTTCT GTCATATCT
13251 GTGAACCTGG GAAGATCACT GTACTCCTTT GTCTGATTTT TTCATGTATA
13301 AAAATTACCT TACAAAGGCT ATTGTGAGGA TGAAATAAGG TAACATATGG
13351 CACATAATAA GTGTTCTGTA TATGCTTCTC TCCTCCCTGG TTCTCTGCTT
13401 CCATATCCAT GTCTCTGGAG TTGCCGTAAT TATTTTTTAA ATAGGCATTT
13451 AAAAAATTAT AAAACAAATA TATGATGATT GTGAAAACT AAAACACTGC
13501 ATAAATATAT AAATTACCAA GAAAAGTTA TGTGAGTCAT CCTCAGAAAT
13551 AACTACTCAT AGGTTTTCCC CTATGCCTAA TTCAACAAAT ACATTGAATA
13601 TTGTTAGTAT TGGATCATCT TATGATACCG ATTTTCAGCT TTCTTTTTAA
13651 ATTTAACAA ATGCCTTGAA TATATTTGCA TGTATTCTT TTTAATGATT
13701 TTTGAGGTTT CCATTACACA AATGTGCCAT AATTTGTTT CAGTATCCTT
13751 ATTGATGAAC AGTTGGATTG TTTCTAATTT TTTACTGTTA TAAAAATGCT
13801 ACAGTAAATA CACTTGCACA GAGATCTTGC AAACAGGCAA CCCATTTTAA
13851 TAAATAAATT CACTGGAGTT ATCAAGGATT TCTGGAATGC AGAAATTTCT
13901 TTAGTAATCT ATCTAACTAT ACTCACCTG ATAATGGATA GTTGGTAAGC
13951 AGATAAGTAA AATTCAGCCA TATCTTATGA TTTGTGTTAA AAAAATTTTT
14001 ATATGTTAAG ACTACAATCT TGGGTAGAAT TTGACAGTAA TATCAAAATT
14051 GTCTCATTCA TTTTACTGGT TTGGAGCCAT ATGCATATTA GCCCCCAA
14101 TCCCAACAAA TAGACCACTT TACATTTGTT TCAAACTCTC AGCCTTATCA
14151 AGGTTTAAAG TATCGAGCAT TTCATAGGAT TGCCTTATAG TTGGTCTAAT
14201 TTAACAACCTG AAATAACCAG GCATAAGCAT AATTAACCTT GGACTCAAGA
14251 AGTTGAGTGG CAGCACCTCA GCTGTGGTTC AAAGCATAGC CACTACTACG
14301 CTCTAAACA ATGGAATAAA GTATAAAGCG GTCTCTCAGT CAAGCCTCAC
14351 ACAGGTAAGA GCGGTGACTT TAAGGGAGTA AGATGAAATA TCGTAACATC
14401 ACCCCAGAAA TAATGCTCTC ACTTTGGTTA CTTTATTTGA TTAGTTGATA
14451 TTTGGCATAA GAGAAATCAC TTGTATTTCT CTATTTAACA ACTCTACATT
14501 TAGAACACTT AATTTTCTCA ATCCCCTAAA AAATTAACAT TTACTGCAGA
14551 TGTTTTTACA TTAACAGATT AATGTCTGGA TCATTCTGAA TTTTGAAGA
14601 CCAACATGT TAACATCACT GACATCACTG AAAACCAGCA ATTAATAGCT
14651 GTAACATTGA ATGGTACCTC ACCAAGCCAG CTAATCAGAA ATATCTCCTG
14701 TGTTCACACT CTGTAAGATT TAGCTTTAGC CAAGGTCTTT GCAAAGATTA
14751 ACCAAATAAT GTGTACAGAA GGTACATCCG CTATTGTAAA AATCATTTCA
14801 CTTTGACAGT ACAGAAGAAG CACCAGCCCT TCTGTTTTAG ATGTAGTCCG
14851 TCCTTTTCAA GCTGTATGAT TGTGGACATG TCAACTTAAC ATCTCGGAGT
14901 TTTTATATCT TCATCAGTGG AATGAGAATA ACAACATATA TCTTGTATC
14951 TCACAGGGTT TTTCAGATGA TCAAATGAAG TAATGTGCAG AACTAACCAA
15001 TGTGGGGAAT TATTATCATC ACTGTTACTT TCATATGAAG TGAAGAAAAT
15051 ATTTTAAAC TCAGTAGTTT AATTTACAAT TTAAGTATGT GTTTTAAAGT
15101 GCCTGTTAGC AAAAATTCAC TAGAAGGATG TAGGACACAC TTAAAGTTTT
15151 CATGTAAAAT TTGTGAGTTC TATTTTAAAC TGAATCTTTT GGCCATGTGT
15201 CAACAAATTA ACGTTATCCT TCACCAATG GGTGGGCTTG AAAAAGGCGT
15251 GATGCATAAA TATTACAGT TGTAGGCAAA ATTGTAATGT TATGTATATG
15301 AATACATATT CATTTTTTCA GGGAGAAGGC TTGTAGATT CATCAAGAAA
15351 TCTTTTCAAA GAGTAGATAA TCATTCATGT ATCACTTACC TAGATGCTCA
15401 TGAAATTTTG CCACTTTATA TAATTCCTTA GTTAGCCAAA AGGAGAGTAA
15451 GATGAAGAGG GGGGAAAAAA AAAACTTCTT TGACAAAGAT GGAGAGAAGC
15501 TGTCACTCTT TGTATTCTTT TATCAATCCA GGAAGCCTTT GGTTTTGACA
15551 ATAAGTGGTC TGAGACTTTG TGTACTCCTC AGATAGGTCC CGGAGGACTA
15601 GATTGGTGCC CATCTGCAGA AAACCAGAGG GGATATATTG ACTCTGCAGA
15651 TCTGCCCTTT GATTCTGCCA TCTCTCAGT GGCCCATGCC TTTTGTGGC
15701 AGACTACTGC CCAAGTTATA GACACTAACA CAGGCACACT GAGTATGGGC

```

FIGURE 3, page 5 of 23



18901 ATGAATCCCTC ATTAGCTCTC TAAGACAAAT TTAAATCAGC TATAAGTTTA  
18951 TGTACTAAAT ATGTCTTCAT GATTAGCAAT ATAGATATAC TTTTTTATTA  
19001 TTATTTTCAT TTTGAAAAGT GATTTTTTTT TGTAAGTTTA AAAAACAAAG  
19051 CTTGGTGTTC TTTCTTTTTT CAGTCGGTCC CGGAGAAAAA TGCAAACGGT  
19101 GTCAAATATT TCCATCACGG GGATGCTTGT CATGTACCTG CTTGCCGCCC  
19151 TCTTTGGTTA CCTAACCTTC TATGGTAGGT CACTCTGAAA GTCATTCTCT  
19201 ATATGCAAAT CCTTGTTAGG CTGGTCCTTG ACCTGGGTAG GTATGATTTT  
19251 TAAAAATTGC CTTCTATAAG CATGCTCTAT AGATGACACA TATTCAATTA  
19301 ATATACTATT TTAGTTTTGT CACTTGACCT GAGGAAATGG GGCCTGATTC  
19351 AGCCTGGCTA ACAAGTTACA AGAATTTGTG AATTAACACC TATTTTATAA  
19401 AAAATATCCC TCAAACAAAA TTATTTTCCT CTAGGGATAG ATGATATTTT  
19451 TCTGGCTAGA CTCCATAGTC CAACTCAGGC TACAAGTGAT GAGAATGAAT  
19501 CCACTTGCAT GTGATAAAGC TCCTTTGATG GAATTATTAA CTGCCACACA  
19551 AATAGCAGGG AAACCTGCCAG GTCCTCAAGT TTGAATTTGC CTCCTCTTTA  
19601 CCAGTCAAGT CAAATCTGGG AGCTTGGGAC TTAGGTAAA ATTTCTGACA  
19651 TATCCCATT TATTGTTTGA TACTAAATGA TTTCTTAAGA AAGAGGACAT  
19701 GACAGAATTT CCTTCAATCT AAGAATGCAC CACCAAAAAA AAGTACTAT  
19751 GGCCACATTA GATTATGCCT GCAACATTTT CTCTCTGGCA TCTTAACAGT  
19801 TCACAAAGGG AGTAGGATTG TACTCCTTCC ATGAAGTGTG GCCACATAAA  
19851 CAGATTTTCAT GGAATCACAT ATTGACCTGG TAGCATATGT TTACATGAAT  
19901 CAGTGTATCA ATATAAATAT ATTTTTGTAT AAACCTCCTT TAAAGTTTTT  
19951 TAACTTAATT TTTTCTTAC TGACTTGGTA AATTGAATTG CATGTATGAC  
20001 AAATTGTGGA GGAAAAGATT CAGGAGTAGG CCACCATTG CTTAGGTTTTT  
20051 TTTTCTATTT GACTAATATT TGACTATTAA CCAAACATGT GCTTTAGATT  
20101 GGGCATTAACT TTTTGGCCGG TTGTGAAATA ATGAATGACG AGGTCAATAC  
20151 TACTGAAGGT ATTTTCACTA CTTTTTGTCT GATCTTGAGG TGAATATCCA  
20201 ACTACGCTTG ATTCCATAGA TATTTTCTTG TTATTTGTGC TTGGAGTCTT  
20251 GAATGAAGGT GTTTTCAAGT AGGGCTGCAT CTTCTCTTA GAGTAGTACC  
20301 CACTGGGAGA CCATCTAAAA ATTATACTAA TTTATCCCTG CACGTTACTT  
20351 ATACTTATTT TAATGAGTTT CATAAGACAA GCAAAAACTT GAAAGAGCCC  
20401 AAAAATATCT GTTTTAGTGT GGTGATGGAG TCATAGTTGT TGAGCTTGAA  
20451 AAAATGGTAG CAATCATTCA TCCTAGAGTT TACACACTGG GTTTGTAACC  
20501 TGCAATCAGG GTGGCTGCAC AGGTAGGGAC AGGGGAGGTG GTAGGCTGGG  
20551 AGAGACAATA TGTGGGGCTT GGGTCTCTCA TCCCCTTCAA CAAGAGCACC  
20601 TTGGTCTCTG TCTGATTTGT AATTGCTTCT GTACAGCGGA GATAGATTTA  
20651 TCACAAATGA AATGAGCTTG AGAGGCTCTT TATTTTGTAT TATACCTTCT  
20701 GCAACGTTAT CAGCTTCAGG ACCTCTTTGT TCATTTGAAT GAAGGTTGCA  
20751 TAGCTAATGA GCTCAGAGGC AAGACCAGAG GTGCCTGGAT TCCCAGGCCT  
20801 AGGTCTTTTC CTCTGTTCTG TGTCTCTCT ATAAAATGTT GCCATAAGTG  
20851 ACCTGTGCTG ATTTGACAAC ACCAAGCGGT TTCATTCTCT TTTTCTGT  
20901 GTAGGAGAAG TTGAAGATGA ATTACTTCAT GCCTACAGCA AAGTGTATAC  
20951 ATTAGACATC CCTCTTCTCA TGTTTCGCC TGGCAGTCCT GTGGCAGTAA  
21001 CACTAATCTG GCCCATTTGC CTCTTCCCAG TAAGTACATA AGACTTTGAT  
21051 GAAAGAAACC TACTTGACCC CATAAATTAG TACATGTGTT CTACCTTCAT  
21101 TTTGATTTAA TTATAGGGTG AGTTTGCAAT TGCAATGCCT GAGGATATTA  
21151 TTTTCTCTATA GCATTTTGTG TCACTTAAAA TTGGCCATTT AATGTGTAGA  
21201 TAGAGCAAGT AGTTTCAGGT GGTATTTTTT TAGTGTAGGA AAAAAATCAT  
21251 AAAACTTATT TTTAACTCA AAGTTGAAAA GTGGAGCTGG AGCTTCTGTC  
21301 TTGTGGATTA GTAAAAGTGA GTAGGAGTTC ATATAACTTT GGAACCTTGA  
21351 AAGCCAAAC CATATTAAGT TTCAAATCTT ATTAATTTT ATCACAGTTT  
21401 TGAAGGCATT TCATTTTTTT TCCAGTTTGT TGTGCTGCAA TAATATACAA  
21451 AAGTTGCCTT TTTTAACTG ATGCCTTGAA GGCTAATGAA AAGGGGATTC  
21501 ATGTTAAGTA AATTATATAC CAGAAAAAAA TTTTTCAAAA AACAGTTATG  
21551 CTATCTATCA CATATCTCTC TCACACATGG CCTCTGCCAG ACTCACACCA  
21601 GGTCAACCCCT CCCTGGCATT TGTATTGGT GTCAGTTTGT TCTGAGATCC  
21651 CAGAGCAGAG CTGGTAGTGA AGATTGGGG TGTGTAGTT AAAACCACCA  
21701 CCTAAGGATA AACACAGGTC TTCACCCCTC TGCCAGCTCC TGTTCATAA  
21751 AACTGAATT TACTCATTCA TTTGAGGGGG AAAAAATAA GTGACACAGT  
21801 AACCAGCACT GTCCTGGACA TAATGTTCCA TACAGGGCTG GCATATGAAG  
21851 ACTATTTCTA TAATGACACT GTGGTCACTT TAAATGCAGC TTGTGTGCTG  
21901 AAATATATTT TGGCACATTC CTTTTTCATG AGTGCATGAA ATCAGATCCG  
21951 TACTACTATG GTGGCTAATA TTTTACTCTT AAATCATGTC TTGCCTCTAA  
22001 TATATCTGAA AGTATTTTCA ATGACATACA CATAGCTTTA GCCTAAAATC

FIGURE 3, page 7 of 23

```

22051 AGCTCCGTCT TGGGTACAAG ACAGAAGACA ACTATAAACA GAAGGTATAC
22101 GATAGGGTAA AATTGCCAGG CAAACAACTT CACTGAGAAA AGGATATCTG
22151 GAGCCCTTCT TTTTATGTGT AAAAAATCA CTCATAAAT TTTGGCACAG
22201 TGTAAGCATT CACATCATTG TAGAATCAAA GCATAAGAAA TCTGTGATGT
22251 GCTTCTGTAT TGCTTTATTC ATATTCATAT AGTGTTTTCA AGCCATGGTT
22301 TTAAGGGATT GCCAGAATTG GCCATCGTCA CACAGACAGC TGGTAACAGT
22351 TCAACTAGTG CAGCTCATAG CCCAACACTG AGGGCTGCAA TTATTGTCAT
22401 GGAAGTAAA AGTCATTTAC TGATGAACAT TTCACCTCAG CATGGAAAAT
22451 CCAAATCTCC CTTAGAAAT TCTTACCCTA TGTGAGAAAT AAAGCACTGA
22501 TATAAATCTG ACCATCAGGA ACAGCAATAG TGTGTAAACA TTAGATGCCA
22551 TTAGAACCAA AATTGACCAT AAGAACCAGA GTTCAGAAAA ATGACTAACT
22601 GCTGTCCTTC ATTATGTATT TCCACTCAAC ATTAGCATTT ATGAAACATT
22651 TTGCACATTA TCCTGTCCTC ACCCTTGCAA TGTTACATTT ATATAATCTG
22701 TGTAAGTGCT CCACTGCCCC ACAGAGTCAT AAGTCCCTGG GACTTGGTGA
22751 TGTGCACAGT GACTGGCACA GAGGGTGAGC TCTGTCGTGC TTGGGAAGAA
22801 AAATGGTCTT CAAATGAATC TTGCCTTGTC TTGAAATGTA TAAACTGCCT
22851 TTTCTAGCAA AAGCATAGAC ACTCTTTCCC TTGGTGACAT GTGCTACGAA
22901 TTCAGCTGGG TTGAGGATCT GGGCTAAATG AACCAAACCT CCCTATACAT
22951 GAAGGATACA CAGAGATGGT GACAGAGAGT GGTCACTTCC GTGAGTGGAT
23001 CTAATCAAG TCCTCTGAAG CTAATTCFAA TTTTTTTTCT TTAATAAAT
23051 GATAAAAGTT GTTATTGGCG CTTTTGCTTG TTTATTTCTG ATAACCTAGG
23101 GCTCAGATTT TCAATGTGTC AAATGCTGAC TCACAGCATG GTTCTCCTGA
23151 CAGTTTATTT CATTAAAGGA ACTCTTCACC AGTAAGTTTA TTTACTTGCC
23201 TTGATATCTC CACACATTAA TAATAAACT AACAAAACCT AATCTGAATT
23251 AAAATCTATC AGCTTTAGGC ATTATTTTGT GTTCTCCTTC TTCAACATG
23301 GTAACGGGC TCTCTTCTT AGGAGCTTGA GAAGATATGA CTGGGGTTTG
23351 TTTTCTCTA CTTCATTTAT TATCTTCTT TTTTCCAATC AGGTTAGTTT
23401 TTTCTTTTTT AGTAAAAGGT GCATAGTAAC TGCTTGTAGT ATTGTTGAA
23451 CAAGTGAATA AATGAAATGA ATTAAGGTAG TGTTTTCAC AGCAGCCCAA
23501 CATTTCTTTC TCTCTTAGTA GTGGGTGGGG TATCAGTTAT GGAATGGCAC
23551 CTCCTTCCAG AGGACTGATC ATGTCATTTT CAGCTTATGC TTCCCTTTAT
23601 GCAGTAAAGT TTCCATATTT CCATAAAGAA CAAGAAACCA AATAATCCTA
23651 ATGGATATAT AATGAACACA CAGATGAAAA TTTCACCTGC CATGCCTTTG
23701 AAAAAAGATC CCTAGCTACT TGTATTTTCAT CTTATAATTA AAATCAGTCT
23751 TTTCACTTAT GTTTTCTTCA GATCTCCTGT TTTGAAGTGT ATATAGATAT
23801 CAACATAGAA ATGCAGCGTA TATTGCTATC AACTGCAGTG GAGCAGTGAT
23851 TCGTAGGTTT TCCAACATCC TTGCCTTAAG CAAACCTGCA AAATCAAAGT
23901 GTGAGCTACG TCTAAACAAT GGGAGAGGCT TTTTTTTTTT TTTTAAGAGT
23951 TAGAACTAAG ACTCTCACTT CCTCCTGTGC CTCCACATTT TTGACCTTCA
24001 CATTGGGCCC CTGCATCAGA ATACAGCACC CCCTAACAGG CTCCTGTTCA
24051 GGACTCTTTC TCTGGAAATA ACAGATGTTG TCTCTAGAGC TGCATAGAAC
24101 CTTAATGGAA TCATTGTGGG TCAGAGGCC TGGATGGTGC TGGGGACCTC
24151 CCTGACCCAC AGCATCTGAC CCACATTTCC AGGTTTCTAG CGACTTGTGT
24201 CAGTAAAGAA AAAGGCACAT AGCTAAGTGG AAGAGCAGAT GAGGCTTGGT
24251 GGGAATCAGC CAGTGGTCTG CCCTAGCAAA GGTAAACAGA ACTGCTGGGG
24301 GCTTTTGGTC CTAGGCTCAC TACTCAGGGA GGCACCTTAA CATGGAATGA
24351 CCAGCAAGTT TCCTTCTCTG TCTTTTCCAC CACCACCACA AGCCTAGTAC
24401 CTCCCTCCCT CTTTGCTCTG TTGCTCTCTT CGGGAATGCA CTGGAACCA
24451 CTTTCAGTTC TGTTTGGAAT TTCTCTATTC CTTATTCAGA AAGAGGAAGA
24501 AGCTTTTGCA TTTACTCCAA CCGTTCTACC TATTATTCCC ATAACTTTC
24551 TGTGATCTCA TATCATTAGG CCAAATGTTA ATCTTTCTGG GAGCCAGGAG
24601 ACTGCTTTCA CATTAGAGG CCCTGGACAT ATAGGACTGC CTCTAACTCA
24651 CTCTAACTCA GCTTATTGAC TTGAATGCAC CTTTTTAACA AGTACTAAA
24701 AAACAACTG TGACTATTCT CTGAAAATGA GCCTATATCT CATACTTATT
24751 TATTCTGTTT AACACTGTGA AACAAATTAA GTCCTCTGGC ACTATGTATA
24801 TACCATAAAA AGCTTATTG TAAGCCTACT AATTGGACCA GTTTTGACAA
24851 TATTGAATAA GCACTAATTG CAGATCATAA TGTAGAATTA TAGGCTGCTG
24901 AGGAAAACAA TATCACACCA TTTGCTTTCC TCAGTTTCCCT TTTCAGAAATG
24951 AGTTTCATAA TGTTCACTAA TCCAATTTTT AAAATCCTTT ACAAAGTTAT
25001 TCTTAACTA TTTCCAGAGA CTATCTGGTT TGTCATTCTA GAAATGAAAT
25051 TGCCTTTTCA GCCTAAACAG ATGGCCTTAA TTTTGGTGG AGTGGTATGA
25101 AAGGAATGTC ACATGAGAAA CTGCAAGCTA TTTAGCTTGA ATTTTTTGTG
25151 ATTCATACAT GTTTCAAAAT ATATTTTACA TTTTCTCTCT TTAAATGAG

```

FIGURE 3, page 8 of 23

25201 TTCCCATCTC TGCACCTTAA GTGACTTCAG AACTAAAATT TTAAAGTGAA  
25251 CATCAATCAC AGCATTTCCA AAAATGTGAA CTCCTAGCTT AACCGAAGTA  
25301 TTCAC TTATT GGAAAGCTGA TAGAGTAATT CCACTAAGTC CAAAAAGTGT  
25351 CCTCTAAAAG ATTCCAAAGA TAAGAGTGTT TTCAACTTTG TCAAGCTGTA  
25401 CAAACACAAA TGTCACTCCC TCCCTCTGCC CACAGGGATC TTTATCCAGT  
25451 TACAGCAGCG TAACTTGAGC AGCTGCTGCA AACTGAGGCT CTCTTGACCC  
25501 TTCGCC TACT TATTTAGCT GCTAAAATAG GGCTGAAATC TGTCAAGGAT  
25551 CCTGAAGGGA AGGATAAGAT TCCTACTATT CAATTTAATT TAAGCTTTTA  
25601 TTCAGTGCCT GCTGTGTGCA CAACACTAAG CTAGAAAGTC TGAGGAATGT  
25651 TTAGATTATT AGGTCCTGTT CCTTGCCCTTT CATAGATTTA CAATCTATTG  
25701 ATAGGGAGAG CTAAAAAGGA GAGAAAGAGG AAGGAGCAAA CATAAAAACG  
25751 TCAAAATTTT AAAATACCAT TTTAAAATTT TATTTTAAAA TGTTAAATAC  
25801 CATGCAAAAT TAAGGAAAAC CTAGATTCAT AAAAATTCCT TTCACAATCT  
25851 TGTGTAAATC AATTCAGTGC TTGCCCTTAA TGTCTCATCC AGTCTGATGA  
25901 GACATGTTTT GTGATCAACA AGGGTTTTAC TATGTTTCTT AATTATGTGT  
25951 CTTGCCCTGTT ATCTCTTTCT GACCGAGATT ATTTTAAACA ATAAATTCCTG  
26001 AAAACTAAGA AAGTGAAAGC ATAAAATATT GTCTTATAAA ATACGCCAAG  
26051 GAAAAAATGA CACTCCATTT CAAATATCAA AAGTTAGCAT CAAGACTGCA  
26101 CAAGATGAAT GTACAGTCAT GTGTGCTTA CAAATGTGGA CATATTCTGA  
26151 GAAATGCATC TTTAGGCAAT TTTGTCATTG TGCAAACACC ATAGATTGTA  
26201 CTTGCAGCCT AATTGGTGGA GCCTACTATA CACTAAGGCT ATATGGCATA  
26251 GCCTAGTACT CTTAGGCTAC AAACCTGTAC AGCATGTTAC TGTACTGAAT  
26301 AGTGGAGGTA CCTGTAACAT AATGGTAAGT ATTTGTGTCT CCAAACGTAG  
26351 AAAAGCTACT GTAAAAATAC AGTATTACAA CCTTAGGGTA TCACTGTCTT  
26401 ATATGTGGTC TGTTGTTGAC CGAAATGACT ATGCTTAATA CCACTGAACCT  
26451 GTACACTTAA AAATGGTTAA GATGGTAAAT TCTATGTTAT GTATGTTTTA  
26501 TAATAATAAA AAAATTGAAA AAAGCATCAA CATCTTTTCT GGGAAAAAAG  
26551 AAAAAGAAAG AAAATGCATT AGAGTGATGA GAATATTTGA AGTAATAGAT  
26601 AAAGTCAAAA ACAAGAAAT GATCTTGCCCT TTGAACTTTC TTGTTTAAGA  
26651 TTCGTACATC AGTGATCACA CTGTTATTTT CCAAACGACC CTTCAGCTGG  
26701 ATACGACATT TCCTGATTGC AGCTGTGCTT ATTGCACTTA ATAATGTTCT  
26751 GGT CATCCTT GTGCCAACTA TAAAATACAT CTTCCGATTTC ATAGGTGAGT  
26801 TTCAGAAAGG CTTCAATTTG GTCAACCCAA ACTCACGCCT CATTAATATGA  
26851 TGGACAGGGA ACCAGTGCTG GGT CATCCAG ATCCCCGTTT TTTCTCAGGC  
26901 TCATGGATTG CCTTTATCCC TGCGAGGCTC TGGTGATTGA GCTGCTCACT  
26951 GTCTCTTCCCT CTTAACTGAC ACTGGGAGCC ACCTTATAGG TCATTTAGTC  
27001 AAGCTGCTTT TTCTGATAGA TGAGGAAACT GACCCCTATA AAAGTCAAGT  
27051 CATATACCTT GGTGTGGACC CAGGATTTGG ACTTAGGTAT TAGCTCCACC  
27101 ATCAGGAAAA GAGGAAGATA GATTTTACCT GCCAGAAGCT CTCTGATACT  
27151 ACGAGTATCA GCTGAACATT GAAAGGTATC TTCAGAGGAA TAGGAGGTTG  
27201 ATTATATAAA GTGTATTATT AGTATTTCCC CATAACTGCA TGGTCTATTA  
27251 ATTTTCTATT TACTCATTGA GGGTTTACTT AAAC TTAA CACAATCTAA  
27301 AACTTTAAAA TAACCATGGG TAGGTCACCT GCAAAGTAAG AGGTGGATAG  
27351 GGTGTGTCAT GAGTTCAGCC ACCTTAGTAT GTATTTATAT TACTAATCCC  
27401 CTGTAAATTT GTGTAAATT CAGCCTTTTG TTGCTTATTA TATGTTGCAT  
27451 ATACTTATGC AGCTTTGATG TTAGGTACAT TTTAATTGTC TCTATAAACA  
27501 TATCTTCTAT GAATAAATAA CCAAGATGAG CTTATGTGAC TTAAGTGTGT  
27551 GTTTT TAGTG CTAAGTATAG GATAGCTTTA TATTTGGTTT ATTTAAAGTG  
27601 TGTGCTGGCA TCTCCTTTGC TAGGAAC TGC TGGGTAGAC ATTGACCTTG  
27651 CCCTGTGTTT GTCTTCTCAG GGGCTTCTTC TGCCACTATG CTGATTTTTA  
27701 TTCTTCCAGC AGTTTTTTTAT CTTAAACTTG TCAAGAAAGA AACTTTTAGG  
27751 TCACCCCAAA AGGTCGGGGT AAGTAAACCT TGCAATTTC CCCATTATTA  
27801 GTTGTCTTTC CAACTACTTA GAATAAACTA GAAAAACAC ATAGTTCAGA  
27851 AAAATGAATC AATGTACAAG AACCAAAAAT CAAAAATGGG CTAGA ACTTT  
27901 CTGGTAGCAG AGAAAGGGGA CATATTTCTG AAAC TCAA GATTCTACTT  
27951 CAAATATCAA ATATCCTGTG TTGAGTCTGT CATAATGTC AAATAGTAGT  
28001 AGCCTTTCCC ACAGACACAT ATGCTTCAGG CAAATAGCAG TGTCCAATAC  
28051 CAAGCTGCTG TTGTGCTATC CGTGGAATAT CATGCAAGAA GGAATTAGGC  
28101 TCCCTAGCGG TGTATGGA TAATTTAAAT ATTTTGGTCA TGGTTGTTAG  
28151 GTTTGCAAA CCAGAGGAAA GATGTTGCTT TTGTTTTCCC TTCCATAGTA  
28201 CCTGTTGTCC CTGGTGTGGA CTAAGATCCA GAACAGAACC ATTCATCGTT  
28251 CTGTTAACCT CTTTAGATAC AAAATACAGT CTTATTAAAT TAGAGAGTAC  
28301 ATATTTCTTT TCCATAAGAC TACTATAGAA ACAAATGCTA GAAATAATTG

FIGURE 3, page 9 of 23



```

28351 TTTTTC CAAT AAGGAAATAT TATCTTTCAC TCCTTAATAA AGTCATGTTA
28401 AGGCTTGAAA AGAATATTTT TACTGAATT ACTCTGAATT TTTACCTTGA
28451 AGTCATTTAC CTTTGGGATG TTCTGGGGAC TTCAGGATAA TTTGGTATCA
28501 AAAGGTCCAC CCAGCAGCTT GCTCCCAAT TTTAACTCTA TGTAAGTCCGT
28551 CTTGCTTGGA TTTTACAGC AGTGTGACCT TGGCAAATTA CTTGTCCTGT
28601 TTGTGACCTA TTTTCAGTTT GACCAATTGT GAAATGAGTA CAATTATCTC
28651 CTAGACCCAT TCTAGTGAAA AATGTTTAGT TGCTGCTTTC TTATATGTAG
28701 GATTAGGAGG TTTAAGTATG TGATAAAATG TAAGGCCTCT TCTGGTGTFA
28751 AAATGCTGAA GTATTTTATA TGTAGGTATG TACATATATC CTTATATATG
28801 TGTGTGTATA TTATATGTAT GCACACACAC ACACACATAT ATACACTTTT
28851 TGTTGCAACA TCTATTAAGC TTTTGGTTTT GTTTGCTTTA TAAAATTAGA
28901 ATCATATCAT ATATGCTATT CTTTTTTAAC CTGCTCTTTT TCACCTAAAA
28951 GATTGTAAGC ATTCTCTAGA TTATGAATC TTTTCTGTCT CCTTGATTTT
29001 TAATAATCAC AGGGTATTCC ATCATCTTGG TGACTAAAT CAATTAACCTA
29051 TTAATCCATT GTTGAACCTG TAGGTTGTAT CTCTCCACTG TATTCCTCTT
29101 CTTTCTTCAA CTAGGATTCT AAATTGACTG ATAGGTTAGG CCTGGGCATC
29151 TGAGATATA AGAATAATAT GGCTCAATAT ATAGATCAGA TTGCCATATT
29201 ATGTAACCAA CTAACAAACA AATTGTACTA AGTATGGTTT CTGTGCTCCT
29251 AACAGAGTCT CTCTGAATTA CAGGCTTTAA TTTTCCTTGT GGTGGAATA
29301 TTCTTCATGA TTGGAAGCAT GGCACCTATT ATAATTGACT GGATTTATGA
29351 TCCTCCAAAT TCCAAGCATC ACTAACACAA GGAAAAATAC TTTCTTTTTT
29401 TATTGGAAT GGTACAAAGT TATACTCAA AAGATATTTG AATTATCTTG
29451 ATTGGAATGT TATTCATAGG AAATAACAGG AAGATTCCAA AGACGTTTAC
29501 CAGTAATATC ACCAGGCACC TGCAGAAGAG GAAAATCACT GTTTTGTCA
29551 AGGATGGTTG TGTATGTGTT TAAAATAAAA CCTGTGGTGC ACATTTCTAC
29601 CCAGGTTTTG CTAGAGCAGT GTGAGATGAT GAAGGTGTAT TTTGCTGCT
29651 TTACGAGCAG AATAAGGGTA ACTGCATGTA ACAATCATCA GATAGTACTC
29701 TTTCCCCTGC CGTCTCCTCA TCCTGCACCC CCTAAAAAAG TACCAAACAT
29751 TTGCATTCTC AGAACATCAA ACAAAAATGC CCTGGTGGCA AAGCTATCAC
29801 CATTTAATGT CTTCTCTCAG TCTTGCACCA AAGTCTCTGG TCTGTTTACT
29851 AACAGAGGCA AAAGGCATGT CTTAGGAAC TTTCTGTTT CTGTAAGGTA
29901 CATGAATGGT CAAACACCAG TCTAGAGCAT CTTATGTCA ACAGCAAAAT
29951 AATATTTTGC CCACCCTGTT TGTGACATTG AGTTGTGACT TCTATATTCA
30001 ATAGATTTTT GTAAATGTTA AAACATCTAT ATTTAAATGT TAAAACACTA
30051 AATATAGAGA GGGGCTTTAT TTCAATCATA GAGCAACAAC AAAAATAATG
30101 CTTATAGCTA AACTGCCTGT TCTAGAAAGC ATCTGCTTTT TCATGTTATT
30151 CCTAAATCCT CTTGTCATAC TTTTGTCAAT GAACAATGCT CTCCTCTCG
30201 TCTTCCATCC TCATTAGAA TTTTLAGAAG ACCACAATCG TGGAGATACA
30251 CTACCCAGTA TTGTTTGATA CATTTTTATT TGATAAACAT TCAGTGCAGG
30301 AAATGTGAT TTGCTATATG TTTATGTATA TAATCTTATT CTGTAGTCAT
30351 CAGAAATGTTA ATGTAAGGTA CATTTGATTT TTATTTTTTA CATGTGTAGT
30401 TTTCTTTCTT CACAGTCAA GCATTTATAT TATTGGGGGT GGGGGCAGGG
30451 AATTAAAGTT GTGGGCTCGA AAATCCATTC ATATGTATCT GTCTACAAAT
30501 GTCTGGGGAT AATTTAAATT TGAAACCTAA GTTATATATA GTTTGGCAAT
30551 GCTCTTCTTC AATATTTACA ATAATAGGAT GATCTACAAG AAAATAAGTT
30601 TCTTTTGTCA AATTTTTATC ATACTAAAGT TGTTCTTTTA ATTTAGCATA
30651 TCTAAATAG GAATTAGTTC AGTTTAGCTC ACACAGGTGT TTGCTGACAT
30701 TCATTGGCCA TTTAATACAG TGTGAGTGG TTCTCCGTAA AAGTATAAGT
30751 GCTAACACTA CGAAGAAATG CACACGATCA TTCTTGCTCA CTTCTATAAC
30801 AAATCTACAT AAAATGGATT TAAAAATTCC TACTCACAGC CTAAACTTC
30851 TGGAGTTCAC TACCTTTTTT TCAAATCATA GTAAGATCAC TTGTGTATTT
30901 TATATTTTAG TAAAGCCAAT TATGAAGTAC AAGTATCATA CACGTACTTT
30951 TGAGCTACTA TTATTTGAAA AAAATCTGCC AAATAGCATC TTTAGGATAT
31001 ATTTACATTT TCACTCATCT AAAAAGTATA CAAAAATAAA AAGTGAAAAA
31051 AGGTATCTTC TGAATGTTCA AGAGCATCCT ATAGTGCCAA ATAATAAAGC
31101 ACCATTTTTT TCTTCATAAC CAGGATTAAA ATTCATATAT ACTGCAGGGC
31151 AGACATACAT ATGATAGCTT GTGCTGATTA ATTTAACCCC ATTTGTAAAC
31201 AGATGAAAAT TTTATTTTCT TATTTCAATTA ATAAGATGGC TCAATGTATT
31251 GGGAGGCTTC TTTTATTATA CAGAAAGTGT ATATTGGTAT ATAATAAATG
31301 AACTTTTCAA ATGACTATGA TGTGATTTTT GATCTATTGT TAAAGAATGT
31351 TGTGTTATTT GTCCATGAAA CAAAATTTAA AATCCAAATA CTGCTTTTCT
31401 TATATTGGTT TATGTTCCAT TTTCAATGTT ACCTTTGACA CATAACTAAC
31451 ATCTATAGCC ATCATCCTGA AAATAATTGC CATCTTATTT TGGCAAAATA

```

FIGURE 3, page 10 of 23



```

31501 GATATTTAAT CCTAAATTAT TATGATGATT ATAATTTTGG CATCACATAT
31551 ATACCACCTA GAATGAATGT GGAAGAAATG AGTCTTTTAT GGTTAGTTTG
31601 AAAGAATCCA TTGAAGATAG AAAATGAGAG AATAGAAGAA ACCTGAGAAT
31651 AGTAAAAATA AGAGCAGAGA AAATATGGGG GCAGGGAAAA CATGTGAGTG
31701 CTAAGGATTG ATTATGAATG AACGATTAGG GGGATTGATG GATCACAGGG
31751 TAAGTATATG CTTAACTTTA TAAGAAACTT CCACATAGTT TTCCACAGTG
31801 TTTCTACCAT TTTCATTTC ACCCGTACTA CCTACAACCT CCCTGACTC
31851 CACAGCCCTG CCAACATTTG GTGTTGTCTT TTGCATTTTA GCCTTTCTAG
31901 TGGGTCTGAA ATGGTAACTC ATTGTGATTT TCATTTCTGC TTCTGTGACA
31951 ACTAATGTTG AAAACTTTTC AAGTGTTTAA TGGTCACTCA TATATCTTCT
32001 TTTGTGAAGT GTGTATTCAA ATCTTTTGCC CATTTTTTAA ATTTAGGTTA
32051 TGTGTTTTTA TTGGGTATTT GTAGAAGCTC TTTAAATATG GATCCATGTC
32101 CAGATTGCCA ATATATTTTC CCAGTCTATG GTATGGTTGC TTATTTTCCT
32151 AAAGGTGTCT TAATTACATC TTTCTGGGGC CAGGTCACCA TAGCTCAAAG
32201 TTTTGCAATT TATGTCTTAA TGAGATAATA TTAATCAGAG TGGTATAGTC
32251 AAAATTAAAT GTTTTGATGT CCTGGGCCCC TATAGGTAGG ACTGGATCAT
32301 CTAACCAAGA TGCAAAAAAA AAAAAACAAA AAAACAAAAA TAGTACTTGG
32351 AAAAATTAT TTTAAATTAA ACA

```

**FEATURES:**

```

Start: 3000
Exon: 3000-3118
Intron: 3119-7452
Exon: 7453-7543
Intron: 7544-8039
Exon: 8040-8155
Intron: 8156-10894
Exon: 10895-10968
Intron: 10969-11437
Exon: 11438-11530
Intron: 11531-16047
Exon: 16048-16129
Intron: 16130-16215
Exon: 16216-16298
Intron: 16299-16408
Exon: 16409-16467
Intron: 16468-17301
Exon: 17302-17577
Intron: 17578-17709
Exon: 17710-17789
Intron: 17790-19073
Exon: 19074-19174
Intron: 19175-20904
Exon: 20905-21029
Intron: 21030-26649
Exon: 26650-26794
Intron: 26795-27670
Exon: 27671-27768
Intron: 27769-29273
Exon: 29274-29372
Stop: 29373

```

**CHROMOSOME MAP POSITION:**

Chromosome 12

**ALLELIC VARIANTS (SNPs):**

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
1386	T	C	Beyond ORF(5')			
2594	T	C	Beyond ORF(5')			
2757	G	T	Beyond ORF(5')			
6107	C	T	Intron			
6392	T	C	Intron			

FIGURE 3, page 11 of 23

9484	C	G	Intron			
10280	A	G	Intron			
10297	G	A	Intron			
10331	G	A	Intron			
10536	T	C	Intron			
11548	T	C	Intron			
11917	G	T	Intron			
12840	T	-	Intron			
12844	A	-	Intron			
12847	T	-	Intron			
13019	C	-	Intron			
13022	A	G	Intron			
13285	G	A	Intron			
14461	G	C	Intron			
15464	-	G	Intron			
15469	-	A	Intron			
15545	T	C	Intron			
16199	T	C	Intron			
16798	T	C	Intron			
18103	C	T	Intron			
18421	A	G	Intron			
18528	G	A	Intron			
18722	T	C	Intron			
18775	C	G	Intron			
18951	T	C	Intron			
18974	T	G	Intron			
19540	A	C	Intron			
19841	G	A	Intron			
20170	A	C	Intron			
20343	T	C	Intron			
20519	G	A	Intron			
20963	T	C	Exon	411	P	P
21840	G	T	Intron			
22783	C	T	Intron			
22787	G	A	Intron			
22825	T	C	Intron			
22967	A	T	Intron			
23248	A	G	Intron			
23764	G	T	Intron			
23765	C	T	Intron			
24432	A	G	Intron			
24538	C	G	Intron			
24693	T	C	Intron			
24819	C	T	Intron			
25743	C	T	Intron			
26044	G	C	Intron			
26555	G	A	Intron			
27886	A	C	Intron			
31884	T	C	Beyond ORF(3')			
32229	T	A	Beyond ORF(3')			

Context:

DNA

Position

1386 ACCCATATGCATGCTTACTTCTATTCTCTCTAGCTTTTAACCTGCTTCTTTTCATCTT  
TTATGTATATACATTTAGGCTGCCTTATATTAATAATAGTTTCATTTTGTTCCTCCTGC  
TTAAAACACTGTGTGCTATTTTAAATTCTGAGAACTGCTTTCTTTATTTCTAGACAA  
TTCTCTGCCATTATCTTTCTGTTTTGTCTCACCTAGTCTCACAAATCTCTATATTTGG  
AATGACTATCAGTGTATATTGAACCTGTAATCTTATTTTTTCCCCATTCCTCTTAACT  
[T,C]  
CTTATTTGTATTTTCTTTTTTTAAATCTCTTCATGCTATAATTTGAGTGATTTCCACAGA  
TCTGTCTTTCAATTTTATAAGTCTTCCTTCAGCTGAGTTTTTTTAAATTTCAATGATTCT

FIGURE 3, page 12 of 23

ATTTTTTCTTTTTTTAAGAATTCCTTTTTTTGACTCTTTTTGCAACAGCCTGTTCTCC  
 TTTTATATTCCTTTATAATGTTTTATTCTGTGAAAGTTATTCTCTATTTTGAATGTTT  
 TCTTTCAAAATGTCTTTCTTTTATTAATTTAATGTAAAAGTCCCTTTTAAATTGCTTTG

2594

CTGAACTTTCTTTTGTTACTATTCTTAACTTTGGCTTCAGGATCCAAGTGCCTAGAAAGT  
 TACTTCCTAAACTTGATCCTCACCTATGTTGCATATTATCAAGCATTGGTGGTGTAAAT  
 TCTTTCATGTCCAATTAAATTAAAGCAGTAATTTCTTTCTAGTTATTGCTAGTAGAGAC  
 ACTGGTAGATTCTGCCTTGGTAGACCTTCCTCTGTCAACAATTTACTTTGTCTTCCTTT  
 CTTTTAAAACATGTATCCCACTCACAAATACCTAAATTTCTTTGAAGACTGCTGCCATGT  
 [T, C]  
 TTAAGATTTCTTTTTTTTCCATAGTACTAGTAAAACCTGCCATTTTCATTATACATAG  
 GCACTCTATAAATATCTGCTAATTTAGCAATTATTAGTAATTTCTTTCTCTCTCCAT  
 TTCTTCTTTCTTGTATTGGGTAAAGGAACATTTTCCAGGATTTGCTTATGTAAAGTTTCA  
 GGAGTTTCTTTCTTCCCTCCCTTTTACAGAGAGCATACAAAATGTAGATGATTATATTC  
 ACTTATTTTCATTAAATAAAATTATAATGATGTATGTTGTGTTCTGTTTGAGAACAGAG

2757

TTATTGCTAGTAGAGACACTGGTAGATTCTGCCTTGGTAGACCTTCCTCTGTCAACAAT  
 TACTTTTGTCTTCCTTTCTTTTAAAACATGTATCCCACTCACAAATACCTAAATTTCTTT  
 GAAGACTGCTGCCATGTTTTAAGATTTCTTTTTTTTTTCCATAGTACTAGTAAAACCTGC  
 CATTTTCATTATACATAGGCACTCTATAAATATCTGCTAATTTAGCAATTATTAGTAAT  
 TCCTTTCTTCTCTTCCATTTCTTCTTTCTTGTATTGGGTAAAGGAACATTTTCCAGGATTT  
 [G, T]  
 CTTATGTAAAGTTTTTCCAGGATTTCTTTTCTTCCCTTTTACAGAGAGCATACAAAAT  
 GTAGATGATTATATTCATTATTTTCAATTTAAATAAAATTATAATGATGTATGTTGTGTT  
 CTGTTTGCAGAACAGAGTGTCTGAACATCAACACAAAGTGAAGAAGCTTAAGCTGAAG  
 GTACAGTATATTATTACTGAAGGGGCTTGTGTGTGGACAGAAAGCGCTGACAGCTC  
 AAATGGATCCCATGGAAGTGAAGATGTCAACATCGAACCAGATGATGAGAGCAGCAGTG

6107

GTTTCGTGTCTGTTTCTATCTACATCTCATACTGTTTTCTATTCTCAAAAAGTAACCCCT  
 GTCATCCTCTTTTCTTCCAGATTATTTTCCAGGATTAGCTTCTGTTATAAAAAATAGCTT  
 GTACAGATCTCCTACAATAATTATTTCTATTTTATTTCTAAGGTTTATTTATTTATTTA  
 TTGAGACAGACAGAGTTTCACTCTTGTGGCCCATGCTGGAGTGAATGGTGAATCTCGG  
 CTCCTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCTCCTGCTTCAGCCTCCTGAGTAG  
 [C, T]  
 TGGGATTACAGGCGCCTGCCACCACACTCGGCTAACTTTTTGTATTTCTAGTAGAGACGA  
 AGTTTCACCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCAAGTTATCCACCCACCT  
 CAGCCTCCCAAAGTGTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTCTAGGATTAT  
 ATTAATAGAACAATCTTCAATTATTTATCTTTCTTTATCTTTCTTTTCATGTAGGAAAT  
 GTCCTAAAATTTTCAAACCTCAATTTGAAAGCACTTTTAAAATCATACATAGTCGAGCA

6392

CAGCCTCCTGAGTAGCTGGGATTACAGGCGCCTGCCACCACACTCGGCTAACTTTTTGTA  
 TTTCTAGTAGAGACGAAGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCA  
 AGTTATCCACCCACCTCAGCCTCCCAAAGTGTGGGATTACAGGCGTGAGCCACTGTGCC  
 TGGCCTCTAGGATTATATTAATAGAACAATCTTCAATTATTTTATCTTTCTTTATCTTTT  
 TTTTCATGTAGGAAATGTCTAAAATTTTCAAACCTCAATTTGAAAGCACTTTTAAAAT  
 [T, C]  
 ATACATAGTCGAGCATTTTATATAAAAACAACTAAAAAGTCTGTGACATTTTGCAGTATA  
 AAAATGCAATGGCAGCAGCAGGCTTATTAATTGAGCCTCTTGGAAATGTGGCTGGTCCCT  
 AGGTCCGTAGCCTCAAAGGCCCTGGCTTGTAACTGCAGGAGCTGACCAGCACAGCTCTAT  
 AACCAAGTTGTACATCTTCTAGCCTGTGTCCAAGAAAACCAGAATCACAACGCTCTGTGG  
 ATAGTGACATCTTAAAGTTTCTTTCCCTCCCACTCTTTTGCCAGTTTCATTGAATTGCT

9484

GCAACATTTATATCACAAATATGTGCTGTTTATGTTCTGAATATCACATATGATTAGTAA  
 TCACACAGCTATTTGAGGGCTAAGCATCAGGACTATAAATATTTGTATTGTGTTAGTGCT  
 TTGATTGAACCTTTTTATGTATAAATTTCTTCAAGCTGAATGGGTTTTATATCAACTTTA  
 CTTTTATATAAGCCATGTTTTGAAATAAACTAGGATTTTAAATAATCTGAATTTTAAATAGC  
 TATGTATGTAGTCATATATTTGTATGCTTTTGTAAATGTGCTTACCTCTAAGACAAAAAA  
 [C, G]  
 CTGCCTTTCTTTATTAATTATACATACCATTAAATGAATTAGGAAGTTACAGATCACTG  
 ATGAATAGAAATAGGAAAAACTTCCCCAATCCCACAGTCATAGATCATCTTCATGAGAG  
 AAGAATGTTCCACTTTTTTAAATGAGGGCCTCATTTTAGGCTTATAAACACTTAGCAGAT  
 GAATTTGGTCAGAACAATTAAATCACTAAACATCATGGGGTGTGTTTGTGTGTCTAAGT  
 AGCCCAGACTGGATTAAAGCTTTCTCTCTTAATTTATAGCAAGTGACACAGTATTTTAAAG

FIGURE 3, page 13 of 23

10280 ATAAGAGTGCACATAGCTACAGGGGTTATAAAATTTATAATTCATGGTCCAAATGTACA  
TTTGTAGTATTGATTTTCATTGGGAATTACCAAGGGATTAGATCAATTGTGGGAAAGTGT  
ATTTTTTAAAAATAAACAAAGATAAAGATTTTTTTTCTGAATTCAGGTAAAAGGCAGCA  
TTGCTCCTCCATTTATTACGTAGATGCTTCTATCAACATTCTTATTTTGTGCTCCAAAT  
CTTGGATTTGGAAAAATACCAATCCGTATAAACATAAAGAAACCATACATGCATGTGGGG  
[A, G]  
TCCTAACACCAGAAATGACTCTGAATGCAAAAAAAAAAAAAAAAAAAGGGAATTTTC  
GTGCCCCATCCTTAGCTTTCTGCTTTCTCTATTATATATGCAACTGCCTGCCCCCTTA  
TCTTACAAAGTACTTCGTAATCTAATGCACAGGATCAGCAGTAATGCAGCTCAGACTGCA  
TGCTTTGCGCTTTGGATTCCGTAGATTTAGATTAAAGGTTAGTCAGGCTATTGAATAGCC  
CTTCAATTCTAAGTGCTGATGTGAATATCATGCAATATGATGTACATATCCCATGTGC

10297 CTACAGGGGTTATAAAATTTATAATTCATGGTCCAAATGTACATTTGTAGTATTGATTTCT  
ATTGGGAATTACCAAGGGATTAGATCAATTGTGGGAAAGTGTATTTTTTAAAAATAAAC  
AAAGATAAAGATTTTTTTTCTGAATTCAGGTAAAAGGCAGCATTTGCTCCTCCATTTATT  
ACGTAGATGCTTCTATCAACATTCTTATTTTTGTGCTCCAAATCTTGGATTTGGAAAAAT  
ACCAATCCGTATAAACATAAAGAAACCATACATGCATGTGGGGATCCTAACACCAGAAAT  
[G, A]  
ACTCTGAATGCAAAAAAAAAAAAAAAAAAAGGGAATTTTCGTGCCCCATCCTTAGCT  
TTCTCTGCTTTCTCTATTATATATGCAACTGCCTGCCCCCTCTATCTTACAAAGTACTTCG  
TAATCTAATGCACAGGATCAGCAGTAATGCAGCTCAGACTGCATGCTTTGCGCTTTGGAT  
TCCTAGATTTAGATTAAAGGTTAGTCAGGCTATTGAATAGCCCTTCAATTCTAAGTGCT  
GATGTGAATATCATGCAATATGATGTACATATCCCATGTGCTGAGTAAGTAGATGTAG

10331 AAATGTACATTTGTAGTATTGATTTTCATTGGGAATTACCAAGGGATTAGATCAATTGTGG  
GGAAAGTGTATTTTTTAAAAATAAACAAAGATAAAGATTTTTTTTCTGAATTCAGGTAA  
AAGGCAGCATTTGCTCCTCCATTTATTACGTAGATGCTTCTATCAACATTCTTATTTTGT  
GCTCCAAATCTTGGATTTGGAAAAATACCAATCCGTATAAACATAAAGAAACCATACATG  
CATGTGGGGATCCTAACACCAGAAATGACTCTGAATGCAAAAAAAAAAAAAAAAAAAAAA  
[G, A]  
GGAATTTTCGTGCCCCATCCTTAGCTTTCTCTGCTTTCTCTATTATATATGCAACTGCCT  
GCCCCCTCTATCTTACAAAGTACTTCGTAATCTAATGCACAGGATCAGCAGTAATGCAGCT  
CAGACTGCATGCTTTGCGCTTTGGATTCCGTAGATTTAGATTAAAGGTTAGTCAGGCTAT  
TGAATAGCCCTTCAATTCTAAGTGCTGATGTGAATATCATGCAATATGATGTACATATT  
CCCATGTGCTGAGTAAGTAGATGTAGCATTGCTAATGTTGCTATACATTTAGCATCTAA

10536 TACCAATCCGTATAAACATAAAGAAACCATACATGCATGTGGGGATCCTAACACCAGAAA  
TGACTCTGAATGCAAAAAAAAAAAAAAAAAAAGGGAATTTTCGTGCCCCATCCTTAG  
CTTTCTCTGCTTTCTCTATTATATATGCAACTGCCTGCCCCCTCTATCTTACAAAGTACTT  
CGTAATCTAATGCACAGGATCAGCAGTAATGCAGCTCAGACTGCATGCTTTGCGCTTTGG  
ATTCTAGATTTAGATTAAAGGTTAGTCAGGCTATTGAATAGCCCTTCAATTCTAAGTG  
[T, C]  
TGATGTGAATATCATGCAATATGATGTACATATCCCATGTGCTGAGTAAGTAGATGTA  
GCATTTGCTAATGTTGCTATACATTTAGCATCTAAGTTATGAACCAGATTCTACCACTGG  
GTAACATTAAAAAAAGTTAGGGACTTCAGGTATGTAATAATAGCAAATCTATTCTTA  
CGACTTTAAAGGTATGTGTAGAGTTCTGAAAAGAATTTCTCAGCCTCCCCCAATCCAC  
ATACTTTTGGAAAGCTGATGATTGAAAAGATTAAATGTGATCCTTTATTGTAACATCTAAC

11548 ACCATTGATTCTGTTTGGAGAACATTTTGATATATTGCTTATTGGTTTTTGGAGTTGCA  
TCTTTTGGGCTTATAATTTCTATATGATGTTTATTACATGTTTGAGACTCCAGCATGGA  
ATTATATGACAAAAATATTTTAGTCATTAAACAATCTCTTTAAACAAGGCTATTTTATCT  
TTGATTGTAGGCTCTTTGATTTATGAAAAATTAGGAGAAAAGGCATTTGGATGGCCGGGA  
AAAAATTGGAGCTTTTGTTCATTACAATGCAGAACATTGGAGGTAAGGGGATATACTTT  
[T, C]  
CAATGGATCCCATAACTTTCTATAGCGTGTCAATAAATAAGAAAACCTTATGGCAATAA  
ACAGGCACCTTTAGATACAGAAAAATTGCTACTTATAGTTCTTAAATTTTAAATGATAGT  
TTCTTAAATAGGTTTGTGTCCTGCTTTAATTAAAAACAGCAATATCTAAGAATGAAATAA  
CATATAAAACCCTGCCAATTGAATTCTAGAATTAAATATAAAATAAAAGCTTTCTTGAT  
TTTTAATGTTATTATAGCATGAATTATTACTCTTAAAAATTGAAGAATTTGTGCTTATAT

11917 TTTAGATACAGAAAAATTGCTACTTATAGTTCTTAAATTTTAAATGATAGTTTCTTAA  
TAGGTTTGTGCTCCTGCTTTAATTAAAAACAGCAATATCTAAGAATGAAATAACATATAAA

FIGURE 3, page 14 of 23

ACCCTGCCAATTGAATTCTAGAATTAAATATAAAATAAAAGCTTTCTTGATTTTAAATG  
TTATTATAGCATGAATTATTACTCTTAAAAATTGAAGAATTTGTGCTTATATCTGTCATT  
GACAAAACAGTTGACGTTTTCTATGTGTGACTGAGTTCGATTTACTAAACTGAAAAGTGG  
[G, T]  
TGTCTGGGGGAACATAGCCAAATGCTGTGGTCCTTGAAACGCAGCCTGCACTGAGCCAGC  
CCACTAGACAGTGTCTCTGGAGTTTACTAAGGCAAAAGTCTGGCTAGGCATCAAATGCA  
CTATAAACCCCGTTTGTGATTCTATGGATTCTTATAAATCCCAGTGAATTATCATTTC  
CAGTGTAGGACCTAGAAATATATATATATATATTTTAAACAATGTTCTCTCGTTGGTGTGT  
TGCCACACAGCTTCATACTGTTTCTGTTGTGCTTTGGCCCTCAGAAGGCATCCAAACCC

12840 GACTATTGCAGTAGTCTTCTAACTGGTCTTCTGGCTTGAGTTTCCCCTGCTCTCAGATA  
AACTCTAATTTGTTCTCCAGATAAACTTTCTCAAATTTGAGTCTGTTTCTACTTTTGTGCG  
TGCATAAAATTTCTCAGCATGCCTTTATTATTTTCAAGGAAAACTTAAACTCATTTGGAC  
TGACACAAGATCTTCGTCTAGTTCTTCTGCTCAATCTTTCTAAACTTTCTAGCAATGCC  
CATATCTATCTATCTTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATC  
[T, -]  
ATCATCTATCAATTTATCCATCATCTATACCCTACATGTCCTGTGTCAAACCATAACAAA  
TTATATTTATTTCCCTAACAGTACTATTTTAAATATTTTAAAAATCATCCATGCCTTCTT  
TTCACAGGCTACTTTCTCCCCTTGACTGTCTCTCAAAGTCTCCAACCTAACACACACG  
CACACACACACACACACACACACACACACACACATTTTCTCTCTCACTCTGCTCAC  
CTGGTCTATTGCTCCTCTAGACTGGTAAATACTAGTTCCTCTGGGCTCTCATGGTCTGT

12844 ATTGCAGTAGTCTTCTAACTGGTCTTCTGGCTTGAGTTTCCCCTGCTCTCAGATAAACT  
CTAATTTGTTCTCCAGATAAACTTTCTCAAATTTGAGTCTGTTTCTACTTTTGTGCGTGA  
TAAATTTCTCAGCATGCCTTTATTATTTTCAAGGAAAACTTAAACTCATTTGGACTGAC  
ACAAGATCTTCGTCTAGTTCTTCTGCTCAATCTTTCTAAACTTTCTAGCAATGCCATA  
TCTATCTATCTTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATC  
[A, -]  
TCTATCAATTTATCCATCATCTATACCCTACATGTCCTGTGTCAAACCATAACAAATTAT  
ATTTATTTCCCTAACAGTACTATTTTAAATATTTTAAAAATCATCCATGCCTTCTTTTCA  
CAGGCTACTTTCTCCCCTTGACTGTCTCTCAAAGTCTCCAACCTAACACACACGCACA  
CACACACACACACACACACACACACACACACATTTTCTCTCTCACTCTGCTCACCTGG  
TCTATTGCTCCTCTAGACTGGTAAATACTAGTTCCTCTGGGCTCTCATGGTCTGTGTTGT

12847 GCAGTAGTCTTCTAACTGGTCTTCTGGCTTGAGTTTCCCCTGCTCTCAGATAAACTCTA  
ATTTGTTCTCCAGATAAACTTTCTCAAATTTGAGTCTGTTTCTACTTTTGTGCGTGATAA  
AATTCCTCAGCATGCCTTTATTATTTTCAAGGAAAACTTAAACTCATTTGGACTGACACA  
AGATCTTCGTCTAGTTCTTCTGCTCAATCTTTCTAAACTTTCTAGCAATGCCATATCT  
ATCTATCTTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATC  
[T, -]  
ATCAATTTATCCATCATCTATACCCTACATGTCCTGTGTCAAACCATAACAAATTATAT  
TATTTCCCTAACAGTACTATTTTAAATATTTTAAAAATCATCCATGCCTTCTTTTACAG  
GCTACTTTCTCCCCTTGACTGTCTCTCAAAGTCTCCAACCTAACACACACGCACACAC  
ACACACACACACACACACACACACACACACATTTTCTCTCTCACTCTGCTCACCTGGTCT  
ATTGCTCCTCTAGACTGGTAAATACTAGTTCCTCTGGGCTCTCATGGTCTGTGTTGTATC

13019 CTGACACAAGATCTTCGTCTAGTTCTTCTGCTCAATCTTTCTAAACTTTCTAGCAATGC  
CCATATCTATCTATCTTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT  
CTATCATCTATCAATTTATCCATCATCTATACCCTACATGTCCTGTGTCAAACCATAACA  
AATTATATTTATTTCCCTAACAGTACTATTTTAAATATTTTAAAAATCATCCATGCCTTC  
TTTTACAGGCTACTTTCTCCCCTTGACTGTCTCTCAAAGTCTCCAACCTAACACACA  
[C, -]  
GCACACACACACACACACACACACACACACACACATTTTCTCTCTCACTCTGCTCA  
CCTGGTCTATTGCTCCTCTAGACTGGTAAATACTAGTTCCTCTGGGCTCTCATGGTCTGT  
TTGTATCTAGTATGTTACTGTTTCTAAAGGATATTTTAAACACTTGAGTAGAGAATA  
AGCTTTTGGAGTCTGATGGACCTGAATTTGAGTCTGTTTCTGTCACTATCTGTGAAGTGG  
GGAAGATCACTGTACTCCTTTGTCTGATTTTTTTCATGTATAAAAATTACCTTACAAAGGC

13022 ACACAAGATCTTCGTCTAGTTCTTCTGCTCAATCTTTCTAAACTTTCTAGCAATGCCCA  
TATCTATCTATCTTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTA  
TCATCTATCAATTTATCCATCATCTATACCCTACATGTCCTGTGTCAAACCATAACAAAT  
TATATTTATTTCCCTAACAGTACTATTTTAAATATTTTAAAAATCATCCATGCCTTCTTT  
TCACAGGCTACTTTCTCCCCTTGACTGTCTCTCAAAGTCTCCAACCTAACACACACGC

FIGURE 3, page 15 of 23

[A, G]  
CACACACACACACACACACACACACACACACATTTTCTCTCTCACTCTGCTCACCT  
GGTCTATTGCTCCTCTAGACTGGTAAATACTAGTTCTCTGGGCTCTCATGGTCCTGTTT  
GTATCTAGTATGTTACTGTTTCTAAAGGATATTTTAAACACTTGAGTAGAGAATAAGC  
TTTTGGAGTCTGATGGACCTGAATTTGAGTCTGTTTCTGTCACTATCTGTGAACCTGGGA  
AGATCACTGTACTCCTTTGTCTGATTTTTTTCATGTATAAAAATTACCTTACAAAGGCTAT

13285 ACTGTCTCTCAAAGTCTCCAACCCTAACACACACGCACACACACACACACACACACA  
CACACACACACACATTTTCTCTCTCACTCTGCTCACCTGGTCTATTGCTCCTCTAGACTG  
GTAAATACTAGTTCTCTGGGCTCTCATGGTCCTGTTTGTATCTAGTATGTTACTGTTT  
CTAAAGGATATTTTAAACACTTGAGTAGAGAATAAGCTTTTGGAGTCTGATGGACCTGA  
ATTTGAGTCTGTTTCTGTCACTATCTGTGAACCTGGGAAGATCACTGTACTCCTTTGTCT  
[G, A]  
ATTTTTTTCATGTATAAAAATTACCTTACAAAGGCTATTGTGAGGATGAAATAAGGTAACA  
TATGGCACAATAAAGTGTTCTGTATATGCTTCTCTCCTCCCTGGTTCTCTGCTTCCATA  
TCCATGTCTCTGGAGTTGCCTGAATTATTTTTTAAATAGGCATTTAAAAAATTATAAAC  
AAATATATGATGATTGTGAAAACTAAACACTGCATAAATATATAAATTACCAAGAAAA  
GTTTATGTCACTCATCCTCAGAAATAACTACTCATAGGTTTTCCCCTATGCCTAATTCAA

14461 TATCGAGCATTTTCATAGGATTGCCTTATAGTTGGTCTAATTTAACAACCTGAAATAACCAG  
GCATAAGCATAATTAACCCCTGGACTCAAGAAGTTGAGTGGCAGCACCTCAGCTGTGGTTC  
AAAGCATAGCCACTACTACGCTTCTAAACAATGAATAAAGTATAAAGCGGTCTCTCAGT  
CAAGCCTCACACAGGTAAGAGGCGTGACTTTAAGGGAGTAAGATGAAATATCGTAACATC  
ACCCAGAAATAATGCTCTCACTTTGGTTACTTTATTTGATTAGTTGATATTTGGCATAA  
[G, C]  
AGAAATCACTTGTATTTCTCTATTTAACAACCTTACATTTAGAACACTTAATTTTCTCAA  
TCCCCTAAAAAATTACATTTACTGCAGATGTTTTCACATTAACAGATTAATGTCTGGAT  
CATTCTGAATTTTTGAAGACCAACATGTTAACATCACTGACATCACTGAAAACAGCAA  
TTAATAGCTGTAACATTGAATGGTACCTCACCAAGCCAGCTAATCAGAAATATCTCCTGT  
GTTCACTCTGTGAAGATTAGCTTTAGCCAAGGTCTTTGCAAAGATTAACCAATAATG

15464 TGAGTTCTATTTTTAACTGAATCTTTTGGCCATGTGTCAACAAATTAACGTTATCCTTCA  
CCAAATGGTGCGGCTTGAAAAAGGCGTGATGCATAAATATTTACAGTTGTAGGCAAAAT  
GTAATGTTATGTATATGAATACATATTCATTTTTTCAGGGAGAAGGCTTGATAGATTTT  
CAAGAAATCTTTCACAAGAGTAGATAATCATTATGTATCACTTACCTAGATGCTCATGA  
AATTTTGCCACTTTATATAATTCCTTAGTTAGCCAAAAGGAGAGTAAGATGAAGAGGGGG  
[-, G]  
AAAAAAAACCTTCTTTGACAAAGATGGAGAGAAGCTGTCTCTCTGTATTCTTTTATC  
AATCCAGGAAGCCTTTGGTTTGTGACAATAAGTGGTCTGAGACTTTGTGTACTCCTCAGAT  
AGGTCCCGGAGGACTAGATTGGTGCCCATCTGCAGAAAACAGAGGGGATATATTGACTC  
TGCAGATCTGCCCTTTGATTCTGCCATCTCTCAGCTGGCCCATGCCTTTTGTGCCAGAC  
TACTGCCAAGTTATAGACACTAACACAGGCACACTGAGTATGGGCTATGTTGATTTATA

15469 TCTATTTTTAACTGAATCTTTTGGCCATGTGTCAACAAATTAACGTTATCCTTCACAAA  
TGGGTGGGCTTGAAAAAGGCGTGATGCATAAATATTTACAGTTGTAGGCAAAATGTAAT  
GTTATGTATATGAATACATATTCATTTTTTCAGGGAGAAGGCTTGATAGATTTTATCAAGA  
AATCTTTCACAAGAGTAGATAATCATTATGTATCACTTACCTAGATGCTCATGAAATTT  
TGCCACTTTATATAATTCCTTAGTTAGCCAAAAGGAGAGTAAGATGAAGAGGGGGAAAA  
[-, A]  
AAAAAATCTTCTTTGACAAAGATGGAGAGAAGCTGTCTCTCTGTATTCTTTTATCAATCC  
AGGAAGCCTTTGGTTTGTGACAATAAGTGGTCTGAGACTTTGTGTACTCCTCAGATAGGTC  
CCGGAGGACTAGATTGGTGCCCATCTGCAGAAAACAGAGGGGATATATTGACTCTGCAG  
ATCTGCCCTTTGATTCTGCCATCTCTCAGCTGGCCCATGCCTTTTGTGCCAGACTACTG  
CCCAAGTTATAGACACTAACACAGGCACACTGAGTATGGGCTATGTTGATTTATACTAA

15545 AGGCGTGATGCATAAATATTTACAGTTGTAGGCAAAATGTAATGTTATGTATATGAATA  
CATATTCATTTTTTTCAGGGAGAAGGCTTGATAGATTTTATCAAGAAATCTTTCACAAGAGT  
AGATAATCATTATGTATCACTTACCTAGATGCTCATGAAATTTTGCCACTTTATATAAT  
TCCTTAGTTAGCCAAAAGGAGAGTAAGATGAAGAGGGGGGAAAAAATCTTCTTGAC  
AAAGATGGAGAGAAGCTGTCTCTCTGTATTCTTTTATCAATCCAGGAAGCCTTTGGTT  
[T, C]  
TGACAATAAGTGGTCTGAGACTTTGTGTACTCCTCAGATAGGTCGGGAGGACTAGATTG  
GTGCCCATCTGCAGAAAACAGAGGGGATATATTGACTCTGCAGATCTGCCCTTTGATTC

FIGURE 3, page 16 of 23

TGCCATCTCTCAGCTGGCCCATGCCTTTTGTGTCAGACTACTGCCCAAGTTATAGACAC  
TAACACAGGCACACTGAGTATGGGCTATGTTGATTTATAACTAATGAGGGCAGAACCTTA  
GAACTGCAGCTTCACTGTAACTTTGGAGCAGGATTTAACACAGAATCAGCCCTGATACT

16199 AGAAGCTTGAAGCAGTGCCAAATACACAATGACTTTTTTTTCCATTTGGGGGATTAGATG  
TTCATCTTACATATCCCAAATGTCATAACTTGCTTGCATGTGACTTCAGTACTGTCCACA  
CCATTAAGCTGTACATTTTCCATTTTAGCAATGTCAAGCTACCTCTTTATCATTAATA  
TGAACCTACCTGAAGTAATCAGAGCATTATGGGACTTGAAGAAAATACTGGGTATGTCTT  
ATGCTCCCTCTGTGACATCAAGTGACTCATTTACTTGGTCTTTTCTGATTCTAATATCC  
[T, C]  
TGTCTCTCACTTCTAGAGAATGGTACCTCAATGGCAACTACCTCATCATATTTGTGTCTG  
TTGGAATTATTCTTCCACTTTTCGCTCCTTAAAAATTTAGGTAAAGATATTTCTAACTGG  
AAATATTTTTATTTTATTTTACATTTAAATAGGTAGCTAATTGTAGATGCCATATTCA  
CCTTCCAAAATGCTTCTTCTAACTTCTAGGTTATCTTGGCTATACCAGTGGATTTTCTCT  
TACCTGCATGGTGTTTTTTGTAGTGTGGTAAGTGATGTGATGACATGATCCTTGCAGGT

16798 GTTGGTTAGCATGAGTTTTTTTTGTGCCTAAATTAGTGTCTCATTTTGTTCAGCAGCTTC  
ACTAATATGAAATAGTTCTTGTATCACAAGTGATTTTCTTGTAGACTAATTTAGAGCAAA  
AAAAGAGCAGCTACGATTTAAAGTAGTTGAGGTAGAATATCAAAGCTACTACTAATGGT  
TTGGTCTAGGCACACTGGTTATATATGGGGAAAAAGGAAAACCTCAAGCAGGAACATGA  
CAATAATCTGGCATTTAGAACAGCAGAGGAGAGTCCCAGATGAGAAACAAGAAGGCTATA  
[T, C]  
CCATATTCACATGAATCAGCCATTCTCTCTTACACATTCCACCCATTAAGAGAGGACAAG  
AACAGTGGGATTAAAGAAGAAATCCTCCTCTTAGGCCCTGACAAAAGAGGGAATTTCT  
TGCATATCATGAATGCCAAATTTATAAAGCATTTCCCAAAGAGGTAAAGGAGAAGGA  
AAAAAGTTTGAAGACCCATGTACCTTAGTTTGAAGAAATAAGGAAATGATCATCTTT  
CTCATGGAAGGGCATGAAAGAGGGTGGGAAGGATTCTTGCAAAATATTGTCTGTAACT

18103 CATTTTAGCATTTCTAATTTGCTTTGAAATTCTGCTCATATGTTCAAAGATTCTTTAACAG  
GAAACACAGTTTATAGCTTCTCTTTCAGAGAAAATATGTACTCCATCCACTCCTCAGTAA  
CATGCTTTAATCAGAAAGGTGGGAATCAGCCCACCACAGCACTACCTTATCTTCTTCTC  
TCCTTTCTCTCCACCATAATGGTTCAGGGGAGGGGTTTCATGGCAGGTGGACAAGGAGTCG  
ATGGTTTAATAATTTTGGCAGGTGTTGGGAATTTAAATTTGAATTTTGTTCGGAAGAAA  
[C, T]  
GATGTCAGCTGGACTAGAAATGAAAACACCCATGACGACCAAACTTATGGTTAGGGGCA  
GCCTCGATAAGCCAGTGATGTCAATTTATAGTCAGCACCTAACCTTGTCTAGAACACATT  
CATTACAAGAGATGTGTCAATATCTGTCCTTTGTTGTCTTATTTGTACAATAGAGTCAC  
GGCTAGAAAATCTTGTCTTCTCCAGCTGATGGTCTATGGTTCATTGTATTCTTTCCCT  
TTGAAGTTGTTGATATTTGCTTGGGAACAAAGGATATGAACTCATTATAGCTGTTTTCCCT

18421 AAATGAAAACACCCATGACGACCAAACTTATGGTTAGGGGCAGCCTCGATAAGCCAGTG  
ATGTCAATTTATAGTCAGCACCTAACCTTGTCTAGAACACATTCAATACAAGAGATGTGT  
CAATATCTGTCCTTTGTTGTCTTATTTGTACAATAGAGTCACTGGCTAGAAAATCTTGT  
TCTTCCAGCTGATGGTCTATGGTTCATTGTATTCTTTCCCTTTGAAGTTGTTGATATT  
TGCTTGGGAACAAAGGATATGAACTCATTATAGCTGTTTTCCCTTTTCCCTTTAAGGGAGG  
[A, G]  
TATTATATAATAATTCTCAACTTCTTTAATCTAGACATCAGTAACCTCAGTCTTCATTCT  
CACTAAATAGCAAACTTTCCCATAAATCTGATTACCTCATAAAAATTTTCAAGACA  
CTTTCAAGTATTTTGATGTCTTTGATTTACTTTGAAAATTACATGTAGCAGTTACTCCAG  
AAGCCTGACAATTGATCTTTGGCAGCCAGGTTCCTTCTAGAATGGTTTTCAGAAGCTTTT  
CAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTTCTTTTCTCAT

18528 ACAAGAGATGTGTCAATATCTGTCCTTTGTTGTCTTATTTGTACAATAGAGTCACTGGCT  
AGAAAATCTGTTTCTTCCAGCTGATGGTCTATGGTTCATTGTATTCTTTTCCCTTTGA  
AGTTGTTGATATTTGCTTGGGAACAAAGGATATGAACTCATTATAGCTGTTTTCCCTTT  
CCTTTAAGGGAGGATATTTATAATAATTCTCAACTTCTTTAATCTAGACATCAGTAACC  
TCAGTCTTCATTCTCACTAAATAGCAAACTTTCCCATAAATCTGATTACCTCATAA  
[G, A]  
AAATTTTCAAGACACTTTCAAGTATTTTGTGATGTTTGTGATTTACTTTGAAAATTACATGTA  
GCAGTTACTCCAGAAGCCTGACAATTGATCTTTGGCAGCCAGGTTCCTTCTAGAATGGT  
TTCAGAAGCTTTTTCAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGG  
TTCTTTTCCCTCATTTAAAGTCATCTCATTTATGAAATGCAAAAGCTTTCTATGTTAGGAGC  
CTGTTTCATCTTTATGTTAATTATATTCTTATTAGTGGGCAAGCTTACTGACCTACGTG

FIGURE 3, page 17 of 23

18722 TATTATATAATAATTCTCAACTTCTTTAATCTAGACATCAGTAACCTCAGTCTTCATTCT  
CACTAAATAGCAAAACTTTCCCATAAATTCTGATTTACCTCATAAAAAATTTCAGAACA  
CTTTCAAGTATTTTGATGTCTTTGATTTACTTTGAAAATTACATGTAGCAGTTACTCCAG  
AAGCCTGACAATTGATCTTTGGCAGCCAGGTTCTTCTAGAATGGTTTTTCAGAAGCTTTT  
CAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTTCTTTTCTCAT  
[T, C]  
TAAAGTCATCTCATTATGAAATGCAAAAGCTTTCTATGTTAGGAGCCTGTTTCATCTTTA  
TGTTAATTATATTCTTATTAGTGGGCAAGCTTACTGACCTACGTGAAATAGACTGTTCC  
TCTTCTAGGGAAATGATTGTTTTTAAGACTGAAGGACTAGTGTTAAGAAAAATGAAAT  
GAATCCTCATTAGTCTCTAAGACAAATTTAAATCAGCTATAAGTTTATGTACTAAATAT  
GTCTTCATGATTAGCAATATAGATATACTTTTTTATTATTATTTTCATTTTGAAAAGTGA

18775 TCATTCTCACTAAATAGCAAAACTTTCCCATAAATTCTGATTTACCTCATAAAAAATTT  
CAGAACACTTTCAAGTATTTTGATGTCTTTGATTTACTTTGAAAATTACATGTAGCAGTT  
ACTCCAGAAGCCTGACAATTGATCTTTGGCAGCCAGGTTCTTCTAGAATGGTTTTTCAGA  
AGCTTTTCAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTTCTTT  
TCCTCATTTAAAGTCATCTCATTATGAAATGCAAAAGCTTTCTATGTTAGGAGCCTGTTT  
[C, G]  
ATCTTTATGTTAATTATATTCTTATTAGTGGGCAAGCTTACTGACCTACGTGAAATAGA  
CTGTTCTCTTCTAGGGAAATGATTGTTTTTAAGACTGAAGGACTAGTGTTAAGAAAAA  
TGGAATGAATCCTCATTAGCTCTCTAAGACAAATTTAAATCAGCTATAAGTTTATGTAC  
TAAATATGTCTTTCATGATTAGCAATATAGATATACTTTTTTATTATTATTTTCATTTTGA  
AAAGTGATTTTTTTTTTGAAGTTAAAAAACAAAGCTTGGTGTCTTTCTTTTCCAGTC

18951 CAGAAGCTTTTCAGGTAGTCTGGACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTT  
CTTTTCTCATTTTAAAGTCATCTCATTATGAAATGCAAAAGCTTTCTATGTTAGGAGCCT  
GTTTCATCTTTATGTTAATTATATTCTTATTAGTGGGCAAGCTTACTGACCTACGTGAA  
ATAGACTGTTCTCTTCTAGGGAAATGATTGTTTTTAAGACTGAAGGACTAGTGTTTAAG  
AAAAATGAAATGAATCCTCATTAGCTCTCTAAGACAAATTTAAATCAGCTATAAGTTTA  
[T, C]  
GTACTAAATATGTCTTCATGATTAGCAATATAGATATACTTTTTTATTATTATTTTCATT  
TTGAAAAGTGATTTTTTTTTTGTAAAGTTTAAAAACAAAGCTTGGTGTCTTTCTTTTCC  
AGTCGGTCCCGGAGAAAAATGCAACGGGTGTCAAATATTTCCATCACGGGGATGCTTGTC  
ATGTACCTGCTTGCCGCCCTCTTTGGTTACCTAACCTTCTATGGTAGGTCACTCTGAAAG  
TCATTCTCTATATGCAATCCTTGTTAGGCTGGTCTTGACCTGGGTAGGTATGATTTTT

18974 ACTCCTGGCAGTAGTACTTTGCTGACTCTACTAGGTTCTTTTCTCATTTAAAGTCATCT  
CATTATGAAATGCAAAAGCTTTCTATGTTAGGAGCCTGTTTCATCTTTATGTTAATTATA  
TTCTTATTCAGTGGGCAAGCTTACTGACCTACGTGAAATAGACTGTTCTCTCTAGGGA  
AATGATTGTTTTTAAGACTGAAGGACTAGTGTTTAAGAAAAATGGAAATGAATCCTCATT  
AGCTCTCTAAGACAAATTTAAATCAGCTATAAGTTTATGTACTAAATATGTCTTCATGAT  
[T, G]  
AGCAATATAGATATACTTTTTTATTATTATTTTCATTTTGAAAAGTGATTTTTTTTTTGT  
AGTTTAAAAACAAAGCTTGGTGTCTTTCTTTTCCAGTCGGTCCCGGAGAAAAATGCA  
AACGGTGTCAAATATTTCCATCACGGGGATGCTTGTCATGTACCTGCTTGCCGCCCTCTT  
TGGTTACCTAACCTTCTATGGTAGGTCACTCTGAAAGTCATTCTCTATATGCAATCCTT  
GTTAGGCTGGTCTTGACCTGGGTAGGTATGATTTTTAAAAATGCCTTCTATAAGCATG

19540 GGTATGATTTTTTAAAAATGCCTTCTATAAGCATGCTCTATAGATGACACATATTCAATT  
AATACTATTTTAGTTTTGTCACTTGACCTGAGGAAATGGGGCCTGATTCAGCCTGGCT  
AACAAGTTACAAGAATTTGTGAATTAAACACCTATTTTATAAAAAATATCCCTCAAACAAA  
ATTATTTTCTCTAGGGATAGATGATATTTCTCTGGCTAGACTCCATAGTCCAACTCAGG  
CTACAAGTGATGAGAATGAATCCACTTGCATGTGATAAAGCTCCTTTGATGGAATTATTA  
[A, C]  
CTGCCACACAAATAGCAGGGAAACTGCCAGGTCTCAAGTTTGAATTTGCCTCCTCTTTA  
CCAGTCAAGTCAAATCTGGGAGCTTGGGACTTTAGGTAAAATTTCTGACATATCCCATT  
TATTTTGTATATACTAAATGATTTCTTAAGAAAGAGGACATGACAGAATTTCTTCAATCT  
AAGAATGCACCACCAAAAAAAGTGACTATGGCCACATTAGATTATGCCTGCAACATTTT  
CTCTCTGGCATCTTAACAGTTCACAAAGGGAGTAGGATTGTACTCTTCCATGAAGTGTG

19841 CTGCCACACAAATAGCAGGGAAACTGCCAGGTCTCAAGTTTGAATTTGCCTCCTCTTTA  
CCAGTCAAGTCAAATCTGGGAGCTTGGGACTTTAGGTAAAATTTCTGACATATCCCATT

FIGURE 3, page 18 of 23



TATTTTGTATACTAAATGATTTCTTAAGAAAGAGGACATGACAGAATTCCTTCAATCT  
AAGAATGCACCACCAAAAAAAGTGACTATGGCCACATTAGATTATGCCTGCAACATTTT  
CTCTCTGGCATCTTAACAGTTCACAAAGGGAGTAGGATTGTACTCCTCCATGAAGTGTG  
[G,A]  
CCACATAAACAGATTTTATGGAATCACATATTGACCTGGTAGCATATGTTTACATGAATC  
AGTGTATCAATATAAATATATTTTGTATAAACCTCCTTTTAAAGTTTAACTTAATTT  
TTTTCTTACTGACTTGGTAAATTGAATTGCATGTATGACAAATTGTGGAGGAAAAGATTC  
AGGAGTAGGCCACCATTGCTTAGGTTTTTTTCTATTTGACTAATATTTGACTATTAAC  
CAAACATGTGCTTTAGATTGGGCATTAACTTTTTGCCGGTTGTGAAATAATGAATGACGA

20170 TATTGACCTGGTAGCATATGTTTACATGAATCAGTGTATCAATATAAATATATTTTGT  
TAAACCTCCTTTTAAAGTTTAACTTAATTTTTTCTTACTGACTTGGTAAATTGAATT  
GCATGTATGACAAATTGTGGAGGAAAAGATTGAGGAGTAGGCCACCATTGCTTAGGTTT  
TTTTCTATTTGACTAATATTTGACTATTAACCAAACATGTGCTTTAGATTGGGCATTAA  
CTTTTGGCCGGTTGTGAAATAATGAATGACGAGGTCAATACTACTGAAGGTATTTTCACT  
[A,C]  
CTTTTGTCTGATCTTGAGGTGAAAATCCAACACGCTTGATTCCATAGATATTTTCTTG  
TTATTTGTGCTTGGAGTCTGAATGAAGGTGTTTTCAAGTAGGGCTGCATCTTCGTCTTA  
GAGTAGTACCCACTGGGAGACCATCTAAAAATTATACTAATTTATCCCTGCACGTTACTT  
ATACTTATTTTAAATGAGTTTCATAAGACAAGCAAAAACCTTGAAAGAGCCCCAAAAATATCT  
GTTTTAGTGTGGTGATGGAGTCATAGTTGTTGAGCTTGAAAAATGGTAGCAATCATTTCA

20343 TAGGTTTTTTTTCTATTTGACTAATATTTGACTATTAACCAAACATGTGCTTTAGATTGG  
GCATTAACCTTTTTGCGGTTGTGAAATAATGAATGACGAGGTCAATACTACTGAAGGTAT  
TTTCACTACTTTTTGTCTGATCTTGAGGTGAAAATCCAACACGCTTGATTCCATAGATA  
TTTTCTTGTATTTGTCTTGGAGTCTGAATGAAGGTGTTTTCAAGTAGGGCTGCATCT  
TCGTCTTAGAGTAGTACCCACTGGGAGACCATCTAAAAATTATACTAATTTATCCCTGCA  
[T,C]  
GTTACTTATACTTATTTTAAATGAGTTTCATAAGACAAGCAAAAACCTTGAAAGAGCCCCAA  
AATATCTGTTTTAGTGTGGTGATGGAGTCATAGTTGTTGAGCTTGAAAAATGGTAGCAA  
TCATTATCTAGAGTTTACACACTGGGTTTGTAACTGCATCAGGAGTGGCTGCACAGG  
TAGGGACAGGGGAGGTGGTAGGCTGGGAGAGACAATATGTGGGCTTGGGTCTCTCATCC  
CCTTCAACAAGAGCACCTTGGTCTCTGTCTGATTGTGAATTGCTTCTGTACAGCGGAGAT

20519 GATATTTTCTTGTATTTGTGCTTGGAGTCTGAATGAAGGTGTTTTCAAGTAGGGCTGC  
ATCTTCGTCTTAGAGTAGTACCCACTGGGAGACCATCTAAAAATTATACTAATTTATCCC  
TGCACGTTACTTTATCTTATTTTAAATGAGTTTCATAAGACAAGCAAAAACCTTGAAAGAGC  
CCAAAAATATCTGTTTTAGTGTGGTGATGGAGTCATAGTTGTTGAGCTTGAAAAATGGT  
AGCAATCATTCATCTAGAGTTTACACACTGGGTTTGTAACCTGCATCAGGAGTGGCTGC  
[G,A]  
CAGGTAGGGACAGGGGAGGTGGTAGGCTGGGAGAGACAATATGTGGGGCTTGGGTCTCTC  
ATCCCTTCAACAAGAGCACCTTGGTCTCTGTCTGATTTGTAATTGCTTCTGTACAGCGG  
AGATAGATTTATCACAATGTAAATGAGCTTGAGAGGCTCTTTATTTTGTATTATACCTTC  
TGCAACGTTATCAGCTTCAGGACCTCTTGTTCATTGTAATGAAGGTGCATAGCTAATG  
AGCTCAGAGGCAAGACCAGAGGTGCCTGGATTCCAGGCCTAGGTCTTTTCTCTGTTCT

20963 TGAGCTTGAGAGGCTCTTTATTTTGTATTATACCTTCTGCAACGTTATCAGCTTCAGGAC  
CTCTTTGTTTCAATGAAGGTTGCATAGCTAATGAGCTCAGAGGCAAGACCAGAGGT  
GCCTGGATTCCCAGGCCTAGGTCTTTTCTCTGTTCTGTGTTCTCTCTATAAAATGTTGC  
CATAAGTGACCTGTGCTGATTTGACAAACCAAGCGGTTTCATTCTCTTTTTCTGTGTGT  
AGGAGAAGTTGAAGATGAATTACTTCATGCCTACAGCAAAGTGTATACATTAGACATCCC  
[T,C]  
CTTCTCATGGTTTCGCTGGCAGTCTTGTGGCAGTAACACTAACTGTGCCCATTTGTCTCT  
TTCCAGTAAGTACATAAGACTTTGATGAAAGAAACCTACTTGACCCATAAATTAGTAC  
ATGTGTTCTACCTTCATTTTGAATTAATATAGGGTGAATTTGCAATTGCAATGCCTGAG  
GATATTATTTTCTATAGCATTTTGAAGTCACTTAAATTTGGCCATTAAATGTGTAGATAG  
AGCAAGTAGTTTCAGGTGGTATTTTATAGTGTAGGAAAAAATCATAAACTTATTTTT

21840 AAACAGTTATGCTATCTATCACATATCTCTCTCACACATGGCCTCTGCCAGACTCACACC  
AGGTCACCCTCCCTGGCATTTGTCTTGGTGTGAGTTGTTCTGAGATCCCAGAGCAGA  
GCTGGTAGTGAAGATTTGGGCTGTGTGAGTTAAACCAACCACTAAGGATAAACACAGGT  
CTTCAACCCTCCTGCCAGCTCCTGTTTCATAAACACTGAATTTACTCATTTGAGGGG  
GAAAAAATAAGTGACACAGTAACCAGCACTGTCTGGACATAATGTTCCATACAGGGCT

FIGURE 3, page 19 of 23

[G, T]  
GCATATGAAGACTATTTCTATAATGACACTGTGGTCACTTTAAATGCAGCTTGTGTGCTG  
AAATATATTTTGGCACATTCCCTTTTTCATGAGTGCATGAAATCAGATCCGTACTIONTATG  
GTGGCTAATATTTTACTCTTAAATCATGTCTTGCCTCTAATATATCTGAAAGTATTTTTCAG  
ATGACATACACATAGCTTTAGCCTAAAATCAGCTCCGTCTTGGGTACAAGACAGAAGACA  
ACTATAAACAGAAGGTATACGATAGGGTAAAATTGCCAGGCAACAACCTTCACTGAGAAA

22783 TGAGAAATAAAGCACTGATATAAATCTGACCATCAGGAACAGCAATAGTGTGTAACATT  
AGATGCCATTAGAACCAAAATTGACCATAAGAACCAGAGTTCAGAAAAATGACTAACTGC  
TGTCCTTCATTATGTATTTCCACTCAACATTAGCATTTATGAAACATTTTGCACATTATC  
CTGTCTCACCCCTTGCAATGTTACATTTATATAATCTGTGTAAGTGCTCCACTGCCCCAC  
AGAGTCATAAGTCCCTGGGACTTGGTGATGTGCACAGTGACTGGCACAGAGGGTGAGCTC  
[C, T]  
GTCGTGCTTGGGAAGAAAAATGGTCTTCAAATGAATCTTGCCTTGTCTTGAAATGTATAA  
ACTGCCCTTTTCTAGCAAAAGCATAGACACTCTTCCCTTGGTGACATGTGCTACGAATTC  
AGCTGGGTTGAGGATCTGGGCTAAATGAACCAACCTCCCTATACATGAAGGATACACAG  
AGATGGTGACAGAGAGTGGTCACTTCCGTGAGTGGATCTCAATCAAGTCCCTCTGAAGCTA  
AATTCATTTTTTTTTCTTTACTAAAATGATAAAAGTTGTTATTGGCGCTTTTGCTTGTTT

22787 AAATAAAGCACTGATATAAATCTGACCATCAGGAACAGCAATAGTGTGTAACATTAGAT  
GCCATTAGAACCAAAATTGACCATAAGAACCAGAGTTCAGAAAAATGACTAACTGCTGTC  
CTTCATTATGTATTTCCACTCAACATTAGCATTTATGAAACATTTTGCACATTATCCTGT  
CCTCACCCCTTGCAATGTTACATTTATATAATCTGTGTAAGTGCTCCACTGCCCCACAGAG  
TCATAAGTCCCTGGGACTTGGTGATGTGCACAGTGACTGGCACAGAGGGTGAGCTCTGTC  
[G, A]  
TGCTTGGGAAGAAAAATGGTCTTCAAATGAATCTTGCCTTGTCTTGAAATGTATAAACTG  
CCTTTTCTAGCAAAAGCATAGACACTCTTCCCTTGGTGACATGTGCTACGAATTCAGCT  
GGGTTGAGGATCTGGGCTAAATGAACCAACCTCCCTATACATGAAGGATACACAGAGAT  
GGTGACAGAGAGTGGTCACTTCCGTGAGTGGATCTCAATCAAGTCCCTCTGAAGCTAAATT  
CAATTTTTTTTTCTTTACTAAAATGATAAAAGTTGTTATTGGCGCTTTTGCTTGTTTATTT

22825 CAATAGTGTGTAACATTAGATGCCATTAGAACCAAAATTGACCATAAGAACCAGAGTTC  
AGAAAAATGACTAACTGCTGTCTTCAATTATGTATTTCCACTCAACATTAGCATTTATGA  
AACATTTTGCACATTATCCTGTCTCCCTCACCCCTTGCAATGTTACATTTATATAATCTGTGTA  
AGTGCTCCACTGCCCCACAGAGTCATAAGTCCCTGGGACTTGGTGATGTGCACAGTGACT  
GGCACAGAGGGTGAGCTCTGTCTGTCTTGGGAAGAAAAATGGTCTTCAAATGAATCTTGC  
[T, C]  
TTGTCTTGAAATGTATAAACTGCCTTTTCTAGCAAAAGCATAGACACTCTTCCCTTGGT  
GACATGTGCTACGAATTCAGCTGGGTTGAGGATCTGGGCTAAATGAACCAACCTCCCTA  
TACATGAAGGATACACAGAGATGGTGACAGAGAGTGGTCACTTCCGTGAGTGGATCTCAA  
TCAAGTCCCTCTGAAGCTAAATTCAATTTTTTTTTCTTTACTAAAATGATAAAAGTTGTTAT  
TGGCGCTTTTGCTTGTTTATTTCTGTATAACTTAGGGCTCAGATTTTCAATGTGTCAAATG

22967 CCTCACCCCTTGCAATGTTACATTTATATAATCTGTGTAAGTGCTCCACTGCCCCACAGAG  
TCATAAGTCCCTGGGACTTGGTGATGTGCACAGTGACTGGCACAGAGGGTGAGCTCTGTC  
GTGCTTGGGAAGAAAAATGGTCTTCAAATGAATCTTGCCTTGTCTTGAAATGTATAAACT  
GCCTTTTCTAGCAAAAGCATAGACACTCTTCCCTTGGTGACATGTGCTACGAATTCAGC  
TGGGTTGAGGATCTGGGCTAAATGAACCAACCTCCCTATACATGAAGGATACACAGAGA  
[A, T]  
GGTGACAGAGAGTGGTCACTTCCGTGAGTGGATCTCAATCAAGTCCCTCTGAAGCTAAATT  
CAATTTTTTTTTCTTTACTAAAATGATAAAAGTTGTTATTGGCGCTTTTGCTTGTTTATTT  
CGTATAACTTAGGGCTCAGATTTTCAATGTGTCAAATGCTGACTCACAGCATGGTTCTCC  
TGACAGTTTTATTTCAATTTAAGGAACCTTTCACCAGTAAGTTTATTTACTTGCCTTGATAT  
CTCCACACATTAATAATAAACTAACAAAACCTAATCTGAATTAATAATCTATCAGCTTTA

23248 CATGAAGGATACACAGAGATGGTGACAGAGAGTGGTCACTTCCGTGAGTGGATCTCAATC  
AAGTCTCTGAAGCTAAATTCAATTTTTTTTTCTTTACTAAAATGATAAAAGTTGTTATTG  
GCGCTTTTGCTTGTTTATTTCTGTATAACTTAGGGCTCAGATTTTCAATGTGTCAAATGCT  
GACTCACAGCATGGTTCTCCTGACAGTTTTATTTCAATTTAAGGAACCTTTCACCAGTAAGT  
TTATTTACTTGCCTTGATATCTCCACACATTAATAATAAACTAACAAAACCTAATCTGA  
[A, G]  
TTAAAATCTATCAGCTTTAGGCATTATTTTGTGTTCTCCTTCTTTCAACATGGTAAGTGG  
GCTCTCTTTCTTAGGAGCTTGAGAAGATATGACTGGGGTTTGTTTTTCTCTACTTCATTT

FIGURE 3, page 20 of 23

ATTATCTTTCTTTTTTCCAATCAGGTAGTTTTTCTCTTTTAGTAAAAGGTGCATAGTA  
 ACTGCTTGTAGTATTTGTTGAACAAGTGAATAAATGAAATGAATTAAGGTAGTGTTCAG  
 CTAGCAGCCCAACATTTCTTTCTCTTAGTAGTGGGTGGGTATCAGTTATGGAATGGCACCTCCTTCCAGAGGACTGATCATGTCATTTTCAG  
 23764 GAAATGAATTAAGGTAGTGTTCCTACTAGCAGCCCAACATTTCTTTCTCTCTTAGTAGTG  
 GGTGGGGTATCAGTTATGGAATGGCACCTCCTTCCAGAGGACTGATCATGTCATTTTCAG  
 CTTATGCTTCCCTTTATGCAGTAAAGTTTCCATATTTCCATAAAGAACAAGAAACCAAAT  
 AATCCTAATGGATATATAATGAACACACAGATGAAAATTTACCTGCCATGCCTTTGAAA  
 AAAGATCCCTAGCTACTTGTATTTTCATCTTATAATTAAAATCAGTCTTTTCACTTATGTT  
 [G, T]  
 TCTTCAGATCTCCTGTTTTGAAGTGTATATAGATATCAACATAGAAATGCAGCGTATATT  
 GCTATCAACTGCAGTGGAGCAGTGATTTCGTAGGTTTTCCAACATCCTTGCCTTAAGCAAA  
 CCTGCAAAATCAAAGTGTGAGCTACGTCTAAACAATGGGAGAGGCTTTTTTTTTTTTTTT  
 AAGAGTTAGAACTAAGACTCTCACTTCCTCCTGTGCCTCCACATTTTGGACCTTCACATT  
 GGGCCCTGCATCAGAATACAGCACCCCTAACAGGCTCCTGTTTCAGGACTCTTCTCTG  
 23765 AAATGAATTAAGGTAGTGTTCCTACTAGCAGCCCAACATTTCTTTCTCTCTTAGTAGTG  
 GTGGGGTATCAGTTATGGAATGGCACCTCCTTCCAGAGGACTGATCATGTCATTTTCAGC  
 TTATGCTTCCCTTTATGCAGTAAAGTTTCCATATTTCCATAAAGAACAAGAAACCAAATA  
 ATCCTAATGGATATATAATGAACACACAGATGAAAATTTACCTGCCATGCCTTTGAAA  
 AAGATCCCTAGCTACTTGTATTTTCATCTTATAATTAAAATCAGTCTTTTCACTTATGTTT  
 [C, T]  
 CTTTCAGATCTCCTGTTTTGAAGTGTATATAGATATCAACATAGAAATGCAGCGTATATTG  
 CTATCAACTGCAGTGGAGCAGTGATTTCGTAGGTTTTCCAACATCCTTGCCTTAAGCAAA  
 CTGCAAAATCAAAGTGTGAGCTACGTCTAAACAATGGGAGAGGCTTTTTTTTTTTTTTTA  
 AGAGTTAGAACTAAGACTCTCACTTCCTCCTGTGCCTCCACATTTTGGACCTTCACATTG  
 GGCCCTGCATCAGAATACAGCACCCCTAACAGGCTCCTGTTTCAGGACTCTTCTCTG  
 24432 GGATGGTGTGGGGACCTCCCTGACCCACAGCATCTGACCCACATTTCCAGGTTCCCTAGC  
 GACTTGTGTGAGTAAAGAAAAAGGCACATAGCTAAGTGAAGAGCAGATGAGGCTTGGTG  
 GGAATCAGCCAGTGGTCTGCCCTAGCAAAGGTAACAGAACTGCTGGGGGCTTTGGTCC  
 TAGGCTCACTACTCAGGAGGCACTTTAACATGGAATGACCAGCAAGTTTCCCTCCTGAT  
 CTTTTCCACCACCACCAAGCCTAGTACCTCCCTCCTCTTTGCTCTGTTGCTCTCTTC  
 [A, G]  
 GGAATGCACCTGGAACACCTTCAGTTCTGTTTGAATTTTCTTATTCCTTATTCAGAAA  
 GAGGAAGAAGCTTTTGCATTTACTCCAACCGTTCTACCTATTATTTCCATAAACTTTCTG  
 TGATCTCATATCATTAGGCCAAATGTTAATCTTTCTGGGAGCCAGGAGACTGCTTTCACA  
 TTCAGAGGCCCTGGACATATAGGACTGCCTCTAACTCACTCTAACTCAGCTTATGACTT  
 GAATGCACCTTTTTAACAAGTGACTAAAAACAACCTGTGACTATTCTCTGAAAATGAGC  
 24538 GATGAGGCTTGGTGGGAATCAGCCAGTGGTCTGCCCTAGCAAAGGTAACAGAACTGCTG  
 GGGGCTTTTGGTCTAGGCTCACTACTCAGGGAGGCACTTTAACATGGAATGACCAGCAA  
 GTTTCCTTCTGATCTTTTCCACCACCACCACAAGCCTAGTACCTCCCTCCCTCTTTGCT  
 CTGTTGCTCTCTTCGGGAATGCACTGGAAACACCTTCAGTTCTGTTTGAATTTTCTTA  
 TTCCTTATTCAGAAAGAGGAAGAAGCTTTGCATTTACTCCAACCGTTCTACCTATTATT  
 [C, G]  
 CCATAAACTTTCTGTGATCTCATATCATTAGGCCAAATGTTAATCTTTCTGGGAGCCAGG  
 AGACTGCTTTTCACATTGAGAGGCCCTGGACATATAGGACTGCCTCTAACTCACTTAAC  
 CAGCTTATTGACTTGAATGCACCTTTTTAACAAGTGACTAAAAACAACCTGTGACTATT  
 CTCTGAAAATGAGCCTATATCTCATACTTATTTATTCTGTTTAACTGTGAAACAAAT  
 AAGTCTCTGGCACTATGTATATACCATAAAAAGCTTATTTGTAAGCCTACTAATTGGAC  
 24693 CCTAGTACCTCCCTCCCTCTTTGCTCTGTTGCTCTCTTCGGGAATGCACTGGAAACCA  
 TTCAGTTCTGTTTGAATTTTCTTATTCCTTATTCAGAAAGAGGAAGAAGCTTTTGCAAT  
 TACTCCAACCGTTCTACCTATTATTTCCATAAACTTTCTGTGATCTCATATCATTAGGCC  
 AAATGTTAATCTTTCTGGGAGCCAGGAGACTGCTTTACATTCAGAGGCCCTGGACATAT  
 AGGACTGCCTCTAACTCACTCTAACTCAGCTTATTGACTTGAATGCACCTTTTTAACAAG  
 [T, C]  
 GACTAAAAACAACCTGTGACTATTCTCTGAAAATGAGCCTATATCTCATACTTATTTAT  
 TCTGTTTAACTGTGAAACAAATTAAGTCTCTGGCACTATGTATATACCATAAAAAGC  
 TTATTTGTAAGCCTACTAATTGGACAGTTTGAACAATATTGAATAAGCACTAATTGCAG  
 ATCATAATGTAGAATTATAGGCTGCTGAGGAAAAACAATATCACACCATTGCTTTCTCA  
 GTTTCCTTTTCAGAATGAGTTTCATAATGTTCACTAATCCAATTTTAAAATCCTTTACA

FIGURE 3, page 21 of 23

24819 AACC GTTCTACCTATTATTCCCATAACTTTCTGTGATCTCATATCATTAGGCCAAATGT  
TAATCTTTCTGGGAGCCAGGAGACTGCTTTACATTAGAGGCCCTGGACATATAGGACT  
GCCTCTAACTCACTCTAACTCAGCTTATTGACTTGAATGCACCTTTTTAACAAGTACTA  
AAAACAAACCTGTGACTATTCTCTGAAAATGAGCCTATATCTCATACTTATTTATTCTGT  
TTAACTGTGAAACAAATTAAGTCCTCTGGCACTATGTATATAACATAAAAAGCTTATT  
[C, T]  
GTAAGCCTACTAATTGGACCAGTTTTGACAATATTGAATAAGCACTAATTGCAGATCATA  
ATGTAGAATTATAGGCTGCTGAGGAAAACAATATCACACCATTTGCTTTCTCAGTTCC  
TTTTCAGAAATGAGTTTCATAATGTTCACTAATCCAATTTTTAAATCCTTTACAAAGTTA  
TTCTTAACTATTTCCAGAGACTATCTGGTTTGTCTTCTAGAAATGAAATTGCCTTTTC  
AGCCTAAACAGATGGCCTTAATTTTGGTGGAGTGGTATGAAAGGAATGTCACATGAGAA

25743 TATCCAGTTACAGCAGCGTAACCTTGAGCAGCTGCTGCAAACTGAGGCTCTCTTGACCCTT  
CGCCTACTTATTTAGCTGCTAAAATAGGGCTGAAATCTGTCAAGGATCCTGAAGGGAAG  
GATAAGATTCTACTATTCAATTTAATTTAAGCTTTTATTAGTGCCTGCTGTGTGCACA  
ACACTAAGCTAGAAAGTCTGAGGAATGTTTAGATTATTAGGTCCTGTTCTTGCCTTTCA  
TAGATTTACAATCTATTGATAGGAGAGCTAAAAGGAGAGAAAGAGGAAGGAGCAAAACA  
[C, T]  
AAAAACGTCAAAATTTTAAATACCATTTTAAATTTTATTTTAAATGTTAAATACCAT  
GCAAAATTAAGGAAAACCTAGATTTCATAAAAATTCCTTTTACAATCTTGTGTAAATCAAT  
TCAGTGTCTGCCCTTAATGTCTCATCCAGTCTGATGAGACATGTTTTGTGATCAACAAGG  
GTTTTACTATGTTTTCTAATTATGTGTCTTGCTGTTATCTCTTTCTGACCGAGATTATT  
TTTAAACAATAAATCTGAAAACCTAAGAAAGTGAAGCATAAAATATTGTCTTATAAAATA

26044 AAAAAACGTCAAAATTTTAAATACCATTTTAAATTTTATTTTAAATGTTAAATACCAT  
GCAAAATTAAGGAAAACCTAGATTTCATAAAAATTCCTTTTACAATCTTGTGTAAATCAAT  
TCAGTGTCTGCCCTTAATGTCTCATCCAGTCTGATGAGACATGTTTTGTGATCAACAAGG  
GTTTTACTATGTTTTCTAATTATGTGTCTTGCTGTTATCTCTTTCTGACCGAGATTATT  
TTTAAACAATAAATCTGAAAACCTAAGAAAGTGAAGCATAAAATATTGTCTTATAAAATA  
[G, C]  
GCCAAGGAAAAATGACACTCCATTTCAAATATCAAAGTTAGCATCAAGACTGCACAAG  
ATGAATGTACAGTCATGTGTTGCTTACAAATGTGGACATATTCTGAGAAATGCATCTTTA  
GGCAATTTTGTCTTGTGCAACACCATAGATTGTAATGTCAGCCTAATTGGTGGAGCCT  
ACTATACACTAAGGCTATATGGCATAGCCTAGTACTCCTAGGCTACAAACCTGTACAGCA  
TGTTACTGTACTGAATAGTGGAGGTACCTGTAACATAATGGTAAGTATTTGTGTCTCCAA

26555 AGTACTCCTAGGCTACAAACCTGTACAGCATGTTACTGTACTGAATAGTGGAGGTACCTG  
TAACATAATGGTAAGTATTTGTGTCTCCAAACGTAGAAAAGCTACTGTAAAATACAGTA  
TTACAACCTTAGGGTATCACTGTCTTATATGTGGTCTGTTGTTGACCGAAATGACTATGC  
TTAATACCACTGAACGTACACTTAAATGTTAAGATGGTAAATCTATGTTATGTAT  
GTTTTATAATAATAAAAAATGAAAAAGCATCAACATCTTTCTGGGAAAAAGAAAA  
[G, A]  
GAAAGAAAAATGCATTAGAGTGATGAGAATATTTGAAGTAATAGATAAAGTCAAAAACAAA  
GAAATGATCTTGCTTTGAACCTTTCTTGTTTAAGATTGCTACATCAGTGATCACACTGTT  
ATTTCCCAAACGACCTTTCAGCTGGATACGACATTTCTGATTGCAGCTGTGCTTATTGC  
ACTTAATAATGTTCTGGTCATCCTTGTGCCAACTATAAAATACATCTTCGGATTTCATAGG  
TGAGTTTCAGAAAGGCTTCAATTTGGTCAACCCAACTCAGCCTCATTAATGATGGAC

27886 GGTTTATTTAAAGTGTGTGCTGGCATCTCCTTTGCTAGGAAGTGTGGGTAAGACATTGA  
CCTTGCCCTGTGTTTGTCTCTCAGGGGCTTCTTCTGCCACTATGCTGATTTTATTCTT  
CCAGCAGTTTTTTATCTTAACTTGTCAAGAAAGAACTTTTAGGTCACCCCAAAGGTC  
GGGGTAAGTAAACCTTGCAATTTCCCCCATTATTAGTTGTTCTTCCAACCTACTAGAATA  
AACTAGAAAATACACATAGTTTCAGAAAAATGAATCAATGTACAAGAACCAAAAATCAAAA  
[A, C]  
TGGGCTAGAACTTTCTGGTAGCAGAGAAAGGGGACATATTTCTGAACTCAAATGATTCT  
ACTTCAAATATCAAATATCCTGTGTTGAGTCTGTCTACATATGTCAAATAGTAGTAGCCTT  
TCCCACAGACACATATGCTTCAGGCAATAGCAGTGTCCAATACCAAGCTGCTGTTGTGC  
TATCCGTGGAAAATCATGCAAGAAGGAATTAGGCTCCCTAGCGGTGTATGGAATAATTT  
AAATATTTTGGTCATGTTGTTAGGTTTGCAAGCCAAAGGAAGATGTTGCTTTTGT

31884 CTTTTATGGTTAGTTTGAAGAATCCATTGAAGATAGAAAATGAGAGAATAGAAGAAACC  
TGAGAATAGTAAAATAAAGAGCAGAGAAAATATGGGGGCAGGGAAAACATGTGAGTGCTA

FIGURE 3, page 22 of 23

AGGATTGATTATGAATGAACGATTAGGGGGATTGATGGATCACAGGGTAAGTATATGCTT  
 AACTTTATAAGAACTTCCACATAGTTTCCACAGTGTTCTACCATTTTCATTTCCACC  
 CGTACTACCTACAACCTCCACTGACTCCACAGCCCTGCCAACATTTGGTGTTGTCTTTTG  
 [T, C]  
 ATTTTAGCCTTTCTAGTGGGTCTGAAATGGTAACTCATTGTGATTTTCATTTCTGCTTCT  
 GTGACAACATAATGTTGAAAACCTTTTCAAGTGTTTAATGGTCACTCATATATCTTCTTTTG  
 TGAAGTGTGTATTCAAATCTTTTGCCCATTTTAAAATTTAGGTTATGTGTTTTATTGG  
 GTATTTGTAGAAGCTCTTTAAATATGGATCCATGTCCAGATTGCCAATATATTTCCCAG  
 TCTATGGTATGGTTGCTTATTTTCCTAAAGGTGTCTTAATTACATCTTCTGGGGCCAGG

32229

TTTCATTTCTGCTTCTGTGACAACATAATGTTGAAAACCTTTTCAAGTGTTTAATGGTCACT  
 CATATATCTTCTTTTGTGAAGTGTGTATTCAAATCTTTTGCCCATTTTAAAATTTAGGT  
 TATGTGTTTTTATTGGGTATTTGTAGAAGCTCTTTAAATATGGATCCATGTCCAGATTGC  
 CAATATATTTTCCCAGTCTATGGTATGGTTGCTTATTTTCCTAAAGGTGTCTTAATTACA  
 TCTTTCTGGGGCCAGGTACCATAGCTCAAAGTTTTGCAATTTATGTCTTAATGAGATAA  
 [T, A]  
 ATTAATCAGAGTGGTATAGTCAAAATTAAATGTTTTGATGTCCTGGGCCCATATAGGTAG  
 GACTGGATCATCTAACCAAGATGCAAAAAAAAAAAAAACAAAAAACAAAAATAGTACTTG  
 GAAAAACTTATTTTAAATTAAACA

TTTCTGCTTCTGTGACAACATAATGTTGAAAACCTTTTCAAGTGTTTAATGGTCACT